

GenCore version 5.1.9  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 21, 2006, 19:13:00 ; Search time 1515 Seconds  
(without alignments)  
11367.296 Million cell updates/sec

Title: US-10-656-598-1

Perfect score: 2470

Sequence: 1 tggggggggaagcgcca.....ataaacatttgtgaatat 2470

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 5244920 seqs, 3486124231 residues

Total number of hits satisfying chosen parameters: 10489840

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N\_Geneseq\_8:\*  
1: Geneseqn1980s:\*  
2: Geneseqn1990s:\*  
3: Geneseqn2000s:\*  
4: Geneseqn2001as:\*  
5: Geneseqn2001bs:\*  
6: Geneseqn2002as:\*  
7: Geneseqn2002bs:\*  
8: Geneseqn2003as:\*  
9: Geneseqn2003bs:\*  
10: Geneseqn2003cs:\*  
11: Geneseqn2003ds:\*  
12: Geneseqn2004as:\*  
13: Geneseqn2004bs:\*  
14: Geneseqn2005s:\*  
15: Geneseqn2006s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2470	100.0	2470	6	ABL65811 Lung canc
2	2470	100.0	2470	6	Aai72250 KIAA0175
3	2470	100.0	2470	6	Abn95227 Gene #172
4	2470	100.0	2470	8	Abx76456 Lung canc
5	2470	100.0	2470	10	Abd80562 Ovarian c
6	2470	100.0	2470	10	Adc38346 Human pro
7	2470	100.0	2470	11	Adm39187 Cancer/an
8	2470	100.0	2470	11	Adm72650 Human TAS
9	2470	100.0	2470	12	Adm72216 Human TAS
10	2470	100.0	2470	12	Adn06035 Antipsori
11	2470	100.0	2470	12	Adl19824 Human sof
12	2470	100.0	2470	13	Adr25674 Breast ca
13	2470	100.0	2470	13	Adu05979 Novel bro
14	2470	100.0	2470	14	Adx07296 Cyclin-de
15	2470	100.0	2470	14	Adw93565 Human mat
16	2470	100.0	2470	14	Ady20529 DNA encod
17	2470	100.0	2470	14	Ady15427 DNA encod
18	2470	100.0	2470	14	Ady61853 Human gen

19	2470	100.0	2470	14	ABE57136
20	2470	100.0	2470	14	ABC82649
21	2470	100.0	2487	13	ADX51545
22	2470	100.0	2960	12	ADQ23801
23	2452	99.3	2463	13	ADX52219
24	2441	98.8	2441	13	ADX52095
25	2438	98.7	2453	8	ABX77595
26	2438	98.7	2501	12	ADP07319
27	2438	98.7	2501	14	ADW93569
28	2438	98.7	2501	15	ABE60929
29	2438	98.7	2501	15	ABE83810
30	2438	98.7	2501	15	ABF71794
31	2438	98.7	2510	14	ADW93566
32	2253	91.2	2368	15	ABF71796
33	2191.4	88.7	2391	13	ACN41843
34	2190.6	88.7	2428	2	ABA92381
35	2159.8	87.4	2519	8	ACA03961
36	2140.4	86.7	2583	13	ACN41839
37	2137.4	86.5	2364	13	ACN41845
38	2127.6	86.1	2162	6	ABK35345
39	2119.4	85.8	2372	13	ACN41844
40	2107.4	85.3	2606	13	ACN41847
41	2041	82.6	2251	15	ABF71798
42	1956.4	79.2	2496	13	ACN41841
43	1955.4	79.2	2530	13	ACN41840
44	1952	79.0	2158	14	ADW93567
45	1896	76.8	2023	13	ADU66638

## ALIGNMENTS

RESULT 1

ABL65811  
ID ABL65811 standard; DNA; 2470 BP.

AC ABL65811;

DT 15-MAY-2002 (first entry)

DE Lung cancer related gene sequence SEQ ID NO:4148.

KW Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;  
KW stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;  
KW cytostatic; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;  
KW gene; ds.

OS Homo sapiens.

PN WO200194629-A2.

PD 13-DEC-2001.

PF 30-MAY-2001; 2001WO-US010838.

XX 05-JUN-2000; 2000US-0209473P.

PR 05-JUN-2000; 2000US-0209531P.

PR 18-SEP-2000; 2000US-0233133P.

PR 18-SEP-2000; 2000US-0233617P.

PR 20-SEP-2000; 2000US-0234009P.

PR 20-SEP-2000; 2000US-0234034P.

PR 22-SEP-2000; 2000US-0234052P.

PR 22-SEP-2000; 2000US-0234509P.

PR 25-SEP-2000; 2000US-0234567P.

PR 25-SEP-2000; 2000US-0234923P.

PR 25-SEP-2000; 2000US-0235077P.

PR 25-SEP-2000; 2000US-0235082P.

PR 25-SEP-2000; 2000US-0235134P.

PR 26-SEP-2000; 2000US-0235280P.

PR 26-SEP-2000; 2000US-0235637P.

PR 26-SEP-2000; 2000US-0235638P.

PR 27-SEP-2000; 2000US-0235711P.

PR 27-SEP-2000; 2000US-0235720P.  
 PR 27-SEP-2000; 2000US-0235840P.  
 PR 27-SEP-2000; 2000US-0235863P.  
 PR 28-SEP-2000; 2000US-0236028P.  
 PR 28-SEP-2000; 2000US-0236032P.  
 PR 28-SEP-2000; 2000US-0236033P.  
 PR 28-SEP-2000; 2000US-0236034P.  
 PR 28-SEP-2000; 2000US-0236109P.  
 PR 28-SEP-2000; 2000US-0236111P.  
 PR 28-SEP-2000; 2000US-0236842P.  
 PR 29-SEP-2000; 2000US-0236891P.  
 PR 02-OCT-2000; 2000US-0237172P.  
 PR 02-OCT-2000; 2000US-0237173P.  
 PR 02-OCT-2000; 2000US-0237278P.  
 PR 02-OCT-2000; 2000US-0237294P.  
 PR 02-OCT-2000; 2000US-0237295P.  
 PR 02-OCT-2000; 2000US-0237316P.  
 PR 02-OCT-2000; 2000US-0237425P.  
 PR 03-OCT-2000; 2000US-0237598P.  
 PR 03-OCT-2000; 2000US-0237604P.  
 PR 03-OCT-2000; 2000US-0237606P.  
 PR 03-OCT-2000; 2000US-0237608P.  
 PR 01-NOV-2000; 2000US-0244867P.  
 PR 01-NOV-2000; 2000US-0245084P.  
 XX  
 PA (AVAL-) AVALON PHARM.  
 XX  
 XX Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;  
 PI Soppet DR, Weaver Z;  
 PI  
 XX WPI; 2002-188264/24.  
 DR  
 XX Screening for anti-neoplastic agent involves exposing cells to a chemical  
 PT agent to be tested for anti-neoplastic activity, and determining a change  
 PT in expression of a gene of a signature gene set.  
 XX  
 PS Claim 1; SEQ ID NO 4148; 44pp: English.  
 XX  
 XX The present invention describes a method (M1) for screening for an anti-  
 CC neoplastic agent. The method involves exposing cells to a chemical agent  
 CC to be tested for anti-neoplastic activity, determining a change in  
 CC expression of at least one gene (I) of a signature gene set, where (I)  
 CC comprises a sequence (S) selected from 8447 sequences (given in ABL61664  
 CC to ABL70110), or is at least 95% identical to (S), where a change in  
 CC expression is indicative of anti-neoplastic activity. (I) has cytostatic  
 CC activity and can be used in gene therapy. M1 can be used for screening an  
 CC anti-neoplastic agent, and can be used for producing a product which is  
 CC the data collected with respect to the anti-neoplastic agent as a result  
 CC of M1, and the data is sufficient to convey the chemical structure and/or  
 CC properties of the agent. M1 can be used in the treatment of cancer such  
 CC as colon, breast, stomach, lung, thyroid, oesophageal, ovarian, kidney,  
 CC prostate or pancreatic cancer, adenocarcinoma, carcinoma, clear cell  
 CC cancer, infiltrating ductal cancer, infiltrating lobular cancer, squamous  
 CC cell carcinoma, neuroendocrine carcinoma, papillary carcinoma and Wilms's  
 CC tumour  
 XX  
 SQ Sequence 2470 BP; 767 A; 488 C; 536 G; 679 T; 0 U; 0 Other;  
 Query Match 100.0%; Score 2470; DB 6; Length 2470;  
 Best Local Similarity 100.0%; Pred. NO. 0;  
 Matches 2470; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 TTGCGGGCGGAGCGGCCACACCCGGCGATCGAAAGATTCTTAGGAACGCGTACCA 60  
 DB 1 TTGCGGGCGGAGCGGCCACACCCGGCGATCGAAAGATTCTTAGGAACGCGTACCA 60  
 QY 61 GCCGCGTCTCTCAGGACAGCAGGCGCCCTGTCTTCTGTCGGGCGCGCTCAGCGTGC 120  
 DB 61 GCCGCGTCTCTCAGGACAGCAGGCGCCCTGTCTTCTGTCGGGCGCGCTCAGCGTGC 120  
 QY 121 TCGCGCCCTCAGGTTCTTTTCTTAATTCGAATAAACTTCGACAGGACTATGAAGATT 180  
 DB 121 TCGCGCCCTCAGGTTCTTTTCTTAATTCGAATAAACTTCGACAGGACTATGAAGATT 180

```
QY 1261 ATAAAAATTATGTGGCGGATTAATAGACTATGATTTGGTGTGAAGATGATTTATCAACAG 1320
Db 1261 ATAAAAATTATGTGGCGGATTAATAGACTATGATTTGGTGTGAAGATGATTTATCAACAG 1320
QY 1321 GTGCTGTACTCTCCCGAAACATACAGTTTACCAAGTACTGGACAGAAATCAAAATGGGGTGG 1380
Db 1321 GTGCTGTACTCTCCCGAAACATACAGTTTACCAAGTACTGGACAGAAATCAAAATGGGGTGG 1380
QY 1381 AATCTAAATCAATTAATCCAGCCTTATGCGAAGACACCTGCAAAATAAATAAGAAACAAAG 1440
Db 1381 AATCTAAATCAATTAATCCAGCCTTATGCGAAGACACCTGCAAAATAAATAAGAAACAAAG 1440
QY 1441 AAAAAATATATCTCTCTAGTCTGCTGTAAGAAATGAAGTACTTTATGTTTCTGAGC 1500
Db 1441 AAAAAATATATCTCTCTAGTCTGCTGTAAGAAATGAAGTACTTTATGTTTCTGAGC 1500
QY 1501 CAAAGACTCCAGTTAATAAGAACACAGCAAGAGAGAAATACTCACTAGCCCAAAATCGTT 1560
Db 1501 CAAAGACTCCAGTTAATAAGAACACAGCAAGAGAGAAATACTCACTAGCCCAAAATCGTT 1560
QY 1561 ACACACTACCCCTCAAAAGCTAGAAACAGCTGCTGAAAGAAATCTCAATTAATAATACCAG 1620
Db 1561 ACACACTACCCCTCAAAAGCTAGAAACAGCTGCTGAAAGAAATCTCAATTAATAATACCAG 1620
QY 1621 TAAATTCACAGGAACAGCAAGTAAATGACAGTGTCTATTAGCCCTGAGAGCGCGTGGC 1680
Db 1621 TAAATTCACAGGAACAGCAAGTAAATGACAGTGTCTATTAGCCCTGAGAGCGCGTGGC 1680
QY 1681 GCTCAGTGAATTTGGATCTCAACCAAGCACATATGAGAGAGACTCCCAAAAGAAAGGAG 1740
Db 1681 GCTCAGTGAATTTGGATCTCAACCAAGCACATATGAGAGAGACTCCCAAAAGAAAGGAG 1740
QY 1741 CCAAAGTGTGGAGCCTTGAAGAGGGTGGATAAGGTTTACCTGTGCTCACCAGGA 1800
Db 1741 CCAAAGTGTGGAGCCTTGAAGAGGGTGGATAAGGTTTACCTGTGCTCACCAGGA 1800
QY 1801 GCAAAAGGAAGGGTCTGCGACAGAGCGGCCACAGAGACTTAAGCTTCACTATATGGA 1860
Db 1801 GCAAAAGGAAGGGTCTGCGACAGAGCGGCCACAGAGACTTAAGCTTCACTATATGGA 1860
QY 1861 CTACAACTAGATTAGTGAATCAGATCAACTGTGTAAGTAAATGCTATCTTCCAA 1920
Db 1861 CTACAACTAGATTAGTGAATCAGATCAACTGTGTAAGTAAATGCTATCTTCCAA 1920
QY 1921 AGAAGCATGTTGACTTTGTACAAAAGGGTTATACACTGAAGTGTCAAAACACAGTCAGATT 1980
Db 1921 AGAAGCATGTTGACTTTGTACAAAAGGGTTATACACTGAAGTGTCAAAACACAGTCAGATT 1980
QY 1981 TTGGGAAAGTGAATGCAATTTGAATTAAGAGTGTGCCAGCTTCAAAACCCGATGG 2040
Db 1981 TTGGGAAAGTGAATGCAATTTGAATTAAGAGTGTGCCAGCTTCAAAACCCGATGG 2040
QY 2041 TGGGTATCAGGAGGCGCGCTTAAGGGCGATGCTGGTGTACAAAGATTAGTGAAG 2100
Db 2041 TGGGTATCAGGAGGCGCGCTTAAGGGCGATGCTGGTGTACAAAGATTAGTGAAG 2100
QY 2101 ACATCCTATCTAGCTGCAAGGTATTAATGATGGATTCTTCCATCTGCGCGGATGAGTGG 2160
Db 2101 ACATCCTATCTAGCTGCAAGGTATTAATGATGGATTCTTCCATCTGCGCGGATGAGTGG 2160
QY 2161 GGTGTGATACACCTACATAAAGACTGTATGATGCTGTTGATTTTAAAGTTCATTGAA 2220
Db 2161 GGTGTGATACACCTACATAAAGACTGTATGATGCTGTTGATTTTAAAGTTCATTGAA 2220
QY 2221 CTACCAACTGTTTCTAAGAGCTATCTTAAGACCAATATCTCTTGTGTTTAAACAAA 2280
Db 2221 CTACCAACTGTTTCTAAGAGCTATCTTAAGACCAATATCTCTTGTGTTTAAACAAA 2280
QY 2281 GATATTATTTGTGTATGAATCTAAATCAAGCCCATCTGTCAATATGTTACTGTCTTTT 2340
Db 2281 GATATTATTTGTGTATGAATCTAAATCAAGCCCATCTGTCAATATGTTACTGTCTTTT 2340
QY 2341 TAATCATGTGTTTGTATATTAATAATTTGTTGACTTCTTTAGATTCACTTCCATATGTG 2400
```

```
Db 2341 TAATCATGTGTTTGTATATTAATAATTTGTTGACTTCTTTAGATTCACTTCCATATGTG 2400
QY 2401 AATGTAAGCTCTTAACATGCTCTTTGTAATGTTGTAATTTCTTCTGAAATAAAACCAT 2460
Db 2401 AATGTAAGCTCTTAACATGCTCTTTGTAATGTTGTAATTTCTTCTGAAATAAAACCAT 2460
QY 2461 TTGTGAATAT 2470
Db 2461 TTGTGAATAT 2470

RESULT 2
AAI72250
ID AAI72250 standard; cDNA; 2470 BP.
XX
AC AAI72250;
XX
DT 02-APR-2002 (first entry)
XX
XX KIAA0175 cDNA.
XX
XX Antisense; KIAA0175; inhibitor; tumour; P21; P53; chemosensitivity;
XX radiosensitivity; gamma-irradiation; hydroxy urea; cell cycle arrest;
XX sensitization; neoplastic disease; chemotherapy; radiotherapy; cancer;
XX ss.
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX CDS 171..2126
XX FT /*tag= a
XX FT /product= "KIAA0175"
XX
XX WO200191739-A2.
XX
XX PD 06-DEC-2001.
XX
XX PF 30-MAY-2001; 2001WO-US017644.
XX
XX PR 31-MAY-2000; 2000US-0208435P.
XX
XX PA (CHIR ) CHIRON CORP.
XX
XX PI Wu B, Seeley TW, Williams LT;
XX
XX DR WPI; 2002-122034/16.
XX
XX P-PSDB; AAB47857.
XX
XX PT New isolated specific inhibitor useful for decreasing the expression of
XX the inhibitor in a mammalian cell.
XX
XX PS Claim 3; Page 56-57; 60pp; English.
XX
XX CC This sequence represents the KIAA0175 cDNA. Antisense oligonucleotides
XX were tested to observe their effect on the kinetics of the KIAA0175
XX transcript. Inhibitors of KIAA0175 may be used for decreasing the
XX expression of KIAA0175 in mammalian cell (preferably tumour cell). They
XX may also be used for decreasing the expression of P21 and P53 in a
XX mammalian cell, for increasing the chemosensitivity and/or
XX radiosensitivity of a mammalian cell by measuring a reduction in gamma-
XX irradiation or hydroxy urea induced P53 or P21 protein levels measuring a
XX reduction in gamma-irradiation or hydroxy urea induced cell cycle arrest
XX and measuring an increase in gamma-irradiation or hydroxy urea induced
XX cell sensitization, and for treating neoplastic disease in a mammal. The
XX inhibitor provides beneficial improvement of chemo and/or radiotherapy
XX despite low transfection efficiency (10 - 70%) and/or transient gene
XX expression. The inhibitor decreases the side effects of the cancer
XX therapy
XX
XX SQ Sequence 2470 BP; 767 A; 488 C; 536 G; 679 T; 0 U; 0 Other;
XX
XX Query Match 100.0%; Score 2470; DB 6; Length 2470;
```

Best Local Similarity 100.0%; Pred. No. 0;				Matches 2470; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
QY	1	TTGGCGGGCGGAGCGGCACACACCGCGGATCGAAAGATCTCTTAGGAACCGCGTACCA	60				
Db	1	TTGGCGGGCGGAGCGGCACACACCGCGGATCGAAAGATCTCTTAGGAACCGCGTACCA	60				
QY	61	GC CGCGCTCTCTCAGGACAGCAGAGCCCGCTGTCTTCTGTGGGCGCGCTCAGCGTGC	120				
Db	61	GC CGCGCTCTCTCAGGACAGCAGAGCCCGCTGTCTTCTGTGGGCGCGCTCAGCGTGC	120				
QY	121	TCGCGCCCTCAGGTTCTTTTCTTAATTCCAAATAAACTTGCAAGAGGACTATGAAAGATT	180				
Db	121	TCGCGCCCTCAGGTTCTTTTCTTAATTCCAAATAAACTTGCAAGAGGACTATGAAAGATT	180				
QY	181	ATGATCAACTTCTCAATATATGAATACATGAATACTATGGACAGGTGGCTTGC	240				
Db	181	ATGATCAACTTCTCAATATATGAATACATGAATACTATGGACAGGTGGCTTGC	240				
QY	241	AGGTCAAACTTGGCTGCCATATCTTACTTGGAGAGATGGTAGCTATAAAATCATG	300				
Db	241	AGGTCAAACTTGGCTGCCATATCTTACTTGGAGAGATGGTAGCTATAAAATCATG	300				
QY	301	AAACACACTAGGAGTGATTTGCCCGGATCAAAACGGAGATGGAGGCTTGAAGACC	360				
Db	301	AAACACACTAGGAGTGATTTGCCCGGATCAAAACGGAGATGGAGGCTTGAAGACC	360				
QY	361	TGAGACATCAGCATATATGTCAACTCTACCATGTGCTAGGACAGCCAAATATTC	420				
Db	361	TGAGACATCAGCATATATGTCAACTCTACCATGTGCTAGGACAGCCAAATATTC	420				
QY	421	TGTTCTTGAAGTCTGCTGAGGAGAGCTGTTTGAATATATATATATATATATAT	480				
Db	421	TGTTCTTGAAGTCTGCTGAGGAGAGCTGTTTGAATATATATATATATATATAT	480				
QY	481	TGTCAGAGAGGAGACCGCGGTTGTCTTCGTCAGATAGTATCTGCTGTTTATGTC	540				
Db	481	TGTCAGAGAGGAGACCGCGGTTGTCTTCGTCAGATAGTATCTGCTGTTTATGTC	540				
QY	541	ACAGCCAGGCTATGCTCAGGAGCTCAAGCCAGCAAAATTTGCTTGTATGATATC	600				
Db	541	ACAGCCAGGCTATGCTCAGGAGCTCAAGCCAGCAAAATTTGCTTGTATGATATC	600				
QY	601	ATAAATTAAGCTGATTTGCTTCTGTGCAAAACCCAAAGGCTTAAACAGGATTAC	660				
Db	601	ATAAATTAAGCTGATTTGCTTCTGTGCAAAACCCAAAGGCTTAAACAGGATTAC	660				
QY	661	ATCTACAGATGCTGTGGAGTCTGGCTTATGACGACCTGAGTTAATACAGGCAAT	720				
Db	661	ATCTACAGATGCTGTGGAGTCTGGCTTATGACGACCTGAGTTAATACAGGCAAT	720				
QY	721	CATATCTTGATCAGAGGAGATGTTTGGAGATGGCATCTGTTATATGTTCTTATGT	780				
Db	721	CATATCTTGATCAGAGGAGATGTTTGGAGATGGCATCTGTTATATGTTCTTATGT	780				
QY	781	GTGGATTTCTACCATTTGATGATGATTAATGCTTATACAAAGAGATTATCAGAG	840				
Db	781	GTGGATTTCTACCATTTGATGATGATTAATGCTTATACAAAGAGATTATCAGAG	840				
QY	841	GAAAAATATGATGTTCCCAAGTGTCTCTCCAGTAGCATCTGCTTCTTCAACAAATGC	900				
Db	841	GAAAAATATGATGTTCCCAAGTGTCTCTCCAGTAGCATCTGCTTCTTCAACAAATGC	900				
QY	901	TGCAGGTGGACCAAGAGACGATTTCTATGAAATCTTATGAAACCATCCCTGGATCA	960				
Db	901	TGCAGGTGGACCAAGAGACGATTTCTATGAAATCTTATGAAACCATCCCTGGATCA	960				
QY	961	TGCAAGATTACAACTATCTCTGTGAGTGGCAAGCAAGATCTTTTATTCACCTCGATG	1020				
Db	961	TGCAAGATTACAACTATCTCTGTGAGTGGCAAGCAAGATCTTTTATTCACCTCGATG	1020				
QY	1021	ATGATTCGCTAAACAGAACTTTCTGTATCATCAAGAAACAAACAGGCAAACTAGGAGTT	1080				
Db	1021	ATGATTCGCTAAACAGAACTTTCTGTATCATCAAGAAACAAACAGGCAAACTAGGAGTT	1080				

Db	1021	ATGATTCGCTAAACAGAACTTTCTGTATCATCAAGAAACAAACAGGCAAACTAGGAGTT	1080
QY	1081	TAATTTCACTGTCGAGTATGATCACCTCACGGCTACCTATCTTCTGCTTCTAGCA	1140
Db	1081	TAATTTCACTGTCGAGTATGATCACCTCACGGCTACCTATCTTCTGCTTCTAGCA	1140
QY	1141	AGGCTCGGGGAAAAACAGTTCGTTTAAAGCTTCTTCTTCTCTCTGTCGACAGCCAC	1200
Db	1141	AGGCTCGGGGAAAAACAGTTCGTTTAAAGCTTCTTCTTCTCTCTGTCGACAGCCAC	1200
QY	1201	CTACCCCATTCACAGACATCAAGTCAAAATTAAGTCTGGAAGATGACCGCAAC	1260
Db	1201	CTACCCCATTCACAGACATCAAGTCAAAATTAAGTCTGGAAGATGACCGCAAC	1260
QY	1261	ATAAAAATTAATCTGCGGAGTTAATAGACTATGATTTGGTGTCAAGATGATTTATCAA	1320
Db	1261	ATAAAAATTAATCTGCGGAGTTAATAGACTATGATTTGGTGTCAAGATGATTTATCAA	1320
QY	1321	GTGCTGCTACTCCCCGAAACATCAAGTTTAAACAGTACTGGAACAGATCAAAATGGG	1380
Db	1321	GTGCTGCTACTCCCCGAAACATCAAGTTTAAACAGTACTGGAACAGATCAAAATGGG	1380
QY	1381	AACTTAATCATTAATCAAGCTTATGAGACACACCTGCAAAATTAATTAAGAAACAA	1440
Db	1381	AACTTAATCATTAATCAAGCTTATGAGACACACCTGCAAAATTAATTAAGAAACAA	1440
QY	1441	AAAAATGATATACTCTTAAGTCTGTAAAGATGAAGAGTACTTTATGTTTCTGAGC	1500
Db	1441	AAAAATGATATACTCTTAAGTCTGTAAAGATGAAGAGTACTTTATGTTTCTGAGC	1500
QY	1501	CAAGACTCCAGTTAATTAAGAAACAGCAATTAAGAGAGAAATCTCACTACGCCAAAT	1560
Db	1501	CAAGACTCCAGTTAATTAAGAAACAGCAATTAAGAGAGAAATCTCACTACGCCAAAT	1560
QY	1561	ACACTACACCTTCAAAAGCTAGAAACAGGCTGCTGAAAGAACTCAATTAATTAAT	1620
Db	1561	ACACTACACCTTCAAAAGCTAGAAACAGGCTGCTGAAAGAACTCAATTAATTAAT	1620
QY	1621	TAAATTAACAGGAAACAGCAAGTTAATGACAGGTTCTATTAGCCCTGAGAGCGGTGC	1680
Db	1621	TAAATTAACAGGAAACAGCAAGTTAATGACAGGTTCTATTAGCCCTGAGAGCGGTGC	1680
QY	1681	GCTCAGTGAAATGGAATCTCAACAGACATATGAGAGAGACTCCCAAAAGAAAGGAG	1740
Db	1681	GCTCAGTGAAATGGAATCTCAACAGACATATGAGAGAGACTCCCAAAAGAAAGGAG	1740
QY	1741	CCAAAGTGTGGGAGCTTGAAGGGGTTGATAAGGTTATCACTGCTCACCAGCA	1800
Db	1741	CCAAAGTGTGGGAGCTTGAAGGGGTTGATAAGGTTATCACTGCTCACCAGCA	1800
QY	1801	GCAAAAGGAAGGTTCTGCCAGAGACGGGCCAGAGACTAAAGCTTCACTATAATGTA	1860
Db	1801	GCAAAAGGAAGGTTCTGCCAGAGACGGGCCAGAGACTAAAGCTTCACTATAATGTA	1860
QY	1861	CTACAACTAGATTAGTGAATCCAGATCAACTGTTGAAATGAATTAATGCTATCTTCCA	1920
Db	1861	CTACAACTAGATTAGTGAATCCAGATCAACTGTTGAAATGAATTAATGCTATCTTCCA	1920
QY	1921	AGAGCATGTTGACCTTGTACAAAGGGTTATACACTGAAGTGTCAACACAGTCAAGTT	1980
Db	1921	AGAGCATGTTGACCTTGTACAAAGGGTTATACACTGAAGTGTCAACACAGTCAAGTT	1980
QY	1981	TTGGGAAAGTGAACAATTCGAATTTGAATTAAGAGTGTCCAGCTTCAAAACCCGATGTG	2040
Db	1981	TTGGGAAAGTGAACAATTCGAATTTGAATTAAGAGTGTCCAGCTTCAAAACCCGATGTG	2040
QY	2041	TGGGATACAGAGGCGCGCTTAAAGGCGATGCTGGGTTTACAAAAGATTAGTGAAG	2100
Db	2041	TGGGATACAGAGGCGCGCTTAAAGGCGATGCTGGGTTTACAAAAGATTAGTGAAG	2100
QY	2101	ACATCCCTATCTAGCTCAAGGTATATTCATGATTTCTTCCATCTCTCCCGGATGAGTGTG	2160
Db	2101	ACATCCCTATCTAGCTCAAGGTATATTCATGATTTCTTCCATCTCTCCCGGATGAGTGTG	2160



QY 2161 GGTGTGATACAGCTACATAAGACCTGTTATGATCGCTTTGATTTAAAGTTCAATGGAA 2220  
DB 2161 GGTGTGATACAGCTACATAAGACCTGTTATGATCGCTTTGATTTAAAGTTCAATGGAA 2220  
QY 2221 CTACCAACTTGTGTTCTTAAGAGCTATCTTAAGACCAATATCTCTTTGTTTAAACAAA 2280  
DB 2221 CTACCAACTTGTGTTCTTAAGAGCTATCTTAAGACCAATATCTCTTTGTTTAAACAAA 2280  
QY 2281 GATATTATTTGTTGTAAGTAATCAAGCCCATCTGTCATTAATGTTACTGTCTTTT 2340  
DB 2281 GATATTATTTGTTGTAAGTAATCAAGCCCATCTGTCATTAATGTTACTGTCTTTT 2340  
QY 2341 TAATCATGTGTTTGTATTAATTAATGTTGTAACCTTCTTGAATTCATTCATATG 2400  
DB 2341 TAATCATGTGTTTGTATTAATTAATGTTGTAACCTTCTTGAATTCATTCATATG 2400  
QY 2401 AATGTAAGCTCTTAAGTAATGTTGTTGTAAGTAATGTTGTTGTTGTTGTTGTTGTT 2460  
DB 2401 AATGTAAGCTCTTAAGTAATGTTGTTGTAAGTAATGTTGTTGTTGTTGTTGTTGTT 2460  
QY 2461 TTGTGAATAT 2470  
DB 2461 TTGTGAATAT 2470

## RESULT 3

ABN95227  
ID ABN95227 standard; DNA; 2470 BP.

AC ABN95227;

DT 13-AUG-2002 (first entry)

DE Gene #1725 used to diagnose liver cancer.

KW Gene; liver cancer; ds; hepatocellular carcinoma; hepatotropic;  
KW metastatic liver tumour; cycostatic; expression profile; disease state;  
KW disease progression; drug toxicity; drug efficacy; drug metabolism.

OS Homo sapiens.

PN WO200229103-A2.

XX 11-APR-2002.

PF 02-OCT-2001; 2001WO-US030589.

PR 02-OCT-2000; 2000US-0237054P.

XX (GENE-) GENE LOGIC INC.

PA Horne D, Alvares C, Peres-Da-Silva S, Vockley JG;  
PI WPI; 2002-426119/45.

DR WPI; 2002-426119/45.

XX Diagnosing and detecting the progression of liver cancer, hepatocellular  
PT carcinoma or metastatic liver tumor in a patient, involves detecting the  
PT level of expression of two or more genes in a liver tissue sample.

PS Claim 1; SEQ ID NO 1725; 298pp; English.

XX The invention relates to a novel method for diagnosing and detecting the  
CC progression of liver cancer, hepatocellular carcinoma or metastatic liver  
CC tumour in a patient, and differentiating metastatic liver cancer from  
CC hepatocellular carcinoma in a patient, involving detecting the level of  
CC expression of two or more genes represented in ABN95227 in a  
CC tissue sample. The method of the invention has hepatotropic, and  
CC cytotostatic activity. The method is useful for diagnosing and detecting  
CC the progression of liver cancer, hepatocellular carcinoma and metastatic  
CC liver carcinoma in a patient. The method is useful for identifying  
CC expression profiles which serve as useful diagnostic markers as well as  
CC markers that can be used to monitor disease states, disease progression,

CC drug toxicity, drug efficacy and drug metabolism. Note: The sequence data  
CC for this patent did not form part of the printed specification, but was  
CC obtained in electronic format directly from WIPO at  
XX ftp.wipo.int/pub/published\_pct\_sequences

SQ Sequence 2470 BP; 767 A; 488 C; 536 G; 679 T; 0 U; 0 Other;

Query Match 100.0%; Score 2470; DB 6; Length 2470;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2470; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTGGGGGGGGAAGCGGCCACACCGCGGATCGAAGAGTTCTTAGGAACCGCGTACCA 60  
DB 1 TTGGGGGGGGAAGCGGCCACACCGCGGATCGAAGAGTTCTTAGGAACCGCGTACCA 60  
QY 61 GCCGGGTCTCTCAGGACAGCAGCGCCCTGCTCTCTGTCGGCGCGCTCAGCGGTGCC 120  
DB 61 GCCGGGTCTCTCAGGACAGCAGCGCCCTGCTCTCTGTCGGCGCGCTCAGCGGTGCC 120  
QY 121 TCCGCCCTCAGGTTCTTTTCTAATTCAAATAAACTTGCAGAGGACTATGAAGATT 180  
DB 121 TCCGCCCTCAGGTTCTTTTCTAATTCAAATAAACTTGCAGAGGACTATGAAGATT 180  
QY 181 ATGATGAACCTTCTCAATATTAATGAATACATGAATTTGGGACAGTGGCTTGC 240  
DB 181 ATGATGAACCTTCTCAATATTAATGAATACATGAATTTGGGACAGTGGCTTGC 240  
QY 241 AGGTCAAACTTCCCTGCCATATCTTACTGGAGAGATGTTAGCTATAAATCATGGATA 300  
DB 241 AGGTCAAACTTCCCTGCCATATCTTACTGGAGAGATGTTAGCTATAAATCATGGATA 300  
QY 301 AAAACACACTAGGAGTGAATTTGCCCGGATCAAAACCGGAGATTGAGGCTTTGAAGAAC 360  
DB 301 AAAACACACTAGGAGTGAATTTGCCCGGATCAAAACCGGAGATTGAGGCTTTGAAGAAC 360  
QY 361 TGAGACATCAGCATATATGTCAACTTCACTACCTAGTCTAGAGACAGCCAAATATTCA 420  
DB 361 TGAGACATCAGCATATATGTCAACTTCACTACCTAGTCTAGAGACAGCCAAATATTCA 420  
QY 421 TGGTCTTGGTACTGCTGCGGAGAGCTGTTTGAATTAATTTCCAGGATCGCC 480  
DB 421 TGGTCTTGGTACTGCTGCGGAGAGCTGTTTGAATTAATTTCCAGGATCGCC 480  
QY 481 TGTCAAGAGAGAGACCCCGGTTGTTCCGTCAGATAGTATCTGCTGTTGTTATGTC 540  
DB 481 TGTCAAGAGAGAGACCCCGGTTGTTCCGTCAGATAGTATCTGCTGTTGTTATGTC 540  
QY 541 ACAGCCAGGCTATGCTCAGAGGACCTCAAGCCAGAAAAATTTGCTGTTGATGATATC 600  
DB 541 ACAGCCAGGCTATGCTCAGAGGACCTCAAGCCAGAAAAATTTGCTGTTGATGATATC 600  
QY 601 ATAAATTAAGCTGATTCACATTTTGGTCTCTGTGCAAAACCCAGGATTAACAGGATTACC 660  
DB 601 ATAAATTAAGCTGATTCACATTTTGGTCTCTGTGCAAAACCCAGGATTAACAGGATTACC 660  
QY 661 ATCTACAGACATGCTGTGGAGTCTGGCTTATGACAGCACCTGAGTTAATACAAGCAAT 720  
DB 661 ATCTACAGACATGCTGTGGAGTCTGGCTTATGACAGCACCTGAGTTAATACAAGCAAT 720  
QY 721 CATATCTTGGATCAGAGCAGATGTTTGGAGCATGGGCATCTGTTATATGTTCTTATGT 780  
DB 721 CATATCTTGGATCAGAGCAGATGTTTGGAGCATGGGCATCTGTTATATGTTCTTATGT 780  
QY 781 GTGATTTCTACCATTTTCATGATGATTAATGCTTTTATACAAGAGATTAATGAGAG 840  
DB 781 GTGATTTCTACCATTTTCATGATGATTAATGCTTTTATACAAGAGATTAATGAGAG 840  
QY 841 GAAATATGATGTTTCCCAAGTGGCTCTCTCCAGTAGCAATCTGCTTCTTCAACAAATGC 900  
DB 841 GAAATATGATGTTTCCCAAGTGGCTCTCTCCAGTAGCAATCTGCTTCTTCAACAAATGC 900  
QY 901 TGCAGGTGACCCAAAGAAACCGGATTTCTATGAAAAATCTATTGAAACCATCTCGGTATCA 960  
DB 901 TGCAGGTGACCCAAAGAAACCGGATTTCTATGAAAAATCTATTGAAACCATCTCGGTATCA 960

Db 901 TGCAGGTGACCCAAAGAAACGGATTCTATGAAAAATCTATTGAACCATCCTCGTATCA 960  
Qy 961 TGCAGATTACAACTACTCTGTGTAGTGCAGCAAGCAAGAAATCCTTTTATTCACCTCGATG 1020  
Db 961 TGCAGATTACAACTACTCTGTGTAGTGCAGCAAGCAAGAAATCCTTTTATTCACCTCGATG 1020  
Qy 1021 ATGATTGCGGTAAACAGAACTTTCGTATACATCACAGAAACAAACAGGCAAAACAAATGGAGATT 1080  
Db 1021 ATGATTGCGGTAAACAGAACTTTCGTATACATCACAGAAACAAACAGGCAAAACAAATGGAGATT 1080  
Qy 1081 TAATTTCACTGTGGCAGTAGATCACTCAACGCTACCTATCTTCTGCTTCTAGCCAAAG 1140  
Db 1081 TAATTTCACTGTGGCAGTAGATCACTCAACGCTACCTATCTTCTGCTTCTAGCCAAAG 1140  
Qy 1141 AGGCTCGGGGAAACAGATTGGTTAAGGCTTCTTCTTCTCTGCTGACAGCCAGTG 1200  
Db 1141 AGGCTCGGGGAAACAGATTGGTTAAGGCTTCTTCTTCTCTGCTGACAGCCAGTG 1200  
Qy 1201 CTACCCCATTTACAGACATCAAGTCAAAATAATTGGAGTCTGGAAAGATGTGACCGCAAGTG 1260  
Db 1201 CTACCCCATTTACAGACATCAAGTCAAAATAATTGGAGTCTGGAAAGATGTGACCGCAAGTG 1260  
Qy 1261 ATAAATAATTATGTGGCGGAAATAATAGACTATGATTGGTGTGAAGATGATTTATCAACAG 1320  
Db 1261 ATAAATAATTATGTGGCGGAAATAATAGACTATGATTGGTGTGAAGATGATTTATCAACAG 1320  
Qy 1321 GTGCTGCTACTCCCGAATCATCACGTTTACCAAGTACTGGACAGAAATCAAAATGGGGTGG 1380  
Db 1321 GTGCTGCTACTCCCGAATCATCACGTTTACCAAGTACTGGACAGAAATCAAAATGGGGTGG 1380  
Qy 1381 AATCTAAATCATTAATCTCAGCCTTATGAGAACACCTGCAAAATAAATTAAGAAACAAAG 1440  
Db 1381 AATCTAAATCATTAATCTCAGCCTTATGAGAACACCTGCAAAATAAATTAAGAAACAAAG 1440  
Qy 1441 AAAATCTATATCTCTTAAGTCTGTCTGTAAGAAATCAAGAGTACTTTATGTTCTCGAGC 1500  
Db 1441 AAAATCTATATCTCTTAAGTCTGTCTGTAAGAAATCAAGAGTACTTTATGTTCTCGAGC 1500  
Qy 1501 CAAAGACTCCAGTTAATAGAACACGATTAAGAGAAATCACTACAGCCCAATCGTT 1560  
Db 1501 CAAAGACTCCAGTTAATAGAACACGATTAAGAGAAATCACTACAGCCCAATCGTT 1560  
Qy 1561 ACATACACCTCCTCAAAAGCTAGAAACAGTCCCTGAAAGAAATCCAAATTAATAATACAG 1620  
Db 1561 ACATACACCTCCTCAAAAGCTAGAAACAGTCCCTGAAAGAAATCCAAATTAATAATACAG 1620  
Qy 1621 TAAATTCACAGGAAACAGAACTTAATAGCAGGTGTCAATAGCCTGAGAGCGGTGCG 1680  
Db 1621 TAAATTCACAGGAAACAGAACTTAATAGCAGGTGTCAATAGCCTGAGAGCGGTGCG 1680  
Qy 1681 GCTCAGTGGAAATGGATCTCAACCAAGCAGATATGGAGGAGCTCCAAAAGAAAGGGAG 1740  
Db 1681 GCTCAGTGGAAATGGATCTCAACCAAGCAGATATGGAGGAGCTCCAAAAGAAAGGGAG 1740  
Qy 1741 CCAAAGTGTGTGGAGCCTTGAAGAGGGGTGGATAAGGTTATCACTGTGCTCACCAGGA 1800  
Db 1741 CCAAAGTGTGTGGAGCCTTGAAGAGGGGTGGATAAGGTTATCACTGTGCTCACCAGGA 1800  
Qy 1801 GCAAAAGGAAGGGTCTGCGAGAGACGGGCCCAAGAGACTAAAGCTTCACTATAATGTGA 1860  
Db 1801 GCAAAAGGAAGGGTCTGCGAGAGACGGGCCCAAGAGACTAAAGCTTCACTATAATGTGA 1860  
Qy 1861 CTACAACTAGATTAGTGAATCCAGATCACTGTTGAATGAATAAATGCTTATCTTCCAA 1920  
Db 1861 CTACAACTAGATTAGTGAATCCAGATCACTGTTGAATGAATAAATGCTTATCTTCCAA 1920  
Qy 1921 AGAAGCATGTTGATCTTGTACAAAGGGTTATACCTGAAGTGTCAAAACACAGTCAAGATT 1980  
Db 1921 AGAAGCATGTTGATCTTGTACAAAGGGTTATACCTGAAGTGTCAAAACACAGTCAAGATT 1980  
Qy 1981 TTGGGAAAGTGAACAATGCAATTTGAATAGAGTGTGCCAGCTTCAAAAACCCGATGTGG 2040  
Db 1981 TTGGGAAAGTGAACAATGCAATTTGAATAGAGTGTGCCAGCTTCAAAAACCCGATGTGG 2040

RESULT 4  
ABX76456

ID ABX76456 standard; DNA; 2470 BP.

XX AC ABX76456;

XX XX 02-APR-2003 (first entry)

XX DE Lung cancer-associated polynucleotide #320.

XX DE Lung cancer-associated polynucleotide; gene; ds; cytostatic; emphysema;  
XX KW antiinflammatory; antiasthmatic; non-small cell lung cancer; atelectasis;  
XX KW small cell lung cancer; benign lesion; precancerous lesion; bronchitis;  
XX KW chronic obstructive pulmonary disease; hypersensitivity pneumonitis;  
XX KW interstitial pulmonary fibrosis; fibrosis; asthma; bronchiectasis.

XX OS Unidentified.

XX XX WO200286443-A2.

XX PN 31-OCT-2002.

XX PF 18-APR-2002; 2002WO-US012476.

XX PR 18-APR-2001; 2001US-0284770P.

XX PR 10-MAY-2001; 2001US-0290492P.

XX PR 09-NOV-2001; 2001US-0339245P.

XX PR 13-NOV-2001; 2001US-0350666P.

XX PR 29-NOV-2001; 2001US-0334370P.

XX PR 12-APR-2002; 2002US-0372246P.

XX PA (BOB-) EOS BIOTECHNOLOGY INC.

XX PI Aziz N, Murray R;

XX XX WPI; 2003-093161/08.

XX DR P-PSDB; ABUS6727.

PT Detecting a lung cancer-associated transcript in a cell from a patient

for treating lung cancer, by contacting a biological sample from the patient with a polynucleotide that exhibits increased or decreased expression in lung cancer.

Claim 22; Page 436; 453pp; English.

The invention relates to a method for detecting a lung cancer-associated transcript in a cell from a patient, comprising contacting a biological sample from the patient with a polynucleotide that selectively hybridises to a sequence that is at least 80 % identical to a gene that exhibits increased or decreased expression in lung cancer samples. Lung cancer-associated polynucleotides and polypeptides are used for identifying associated polynucleotides and polypeptides are used for identifying a compound that modulates a lung cancer-associated polypeptide, for inhibiting proliferation of a lung cancer-associated cell to treat lung cancer in a patient and for treating a mammal having lung cancer by administering a modulatory compound identified. The methods are useful for treating lung cancer, such as small cell lung cancer, non-small cell lung cancer or other benign or precancerous lesions, e.g. atelectasis, emphysema, bronchitis, chronic obstructive pulmonary disease, fibrosis, hypersensitivity pneumonitis, interstitial pulmonary fibrosis, asthma and bronchiectasis. The genes, polynucleotides and polypeptides are useful for diagnostic purposes and as targets for screening for therapeutic compounds that modulate lung cancer, such as antibodies. Sequences ABX76124-ABX76474 represent lung cancer-associated polynucleotides of the invention.

Sequence 2470 BP; 767 A; 488 C; 536 G; 679 T; 0 U; 0 Other;

```
Query Match      100.0%; Score 2470; DB 8; Length 2470;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2470; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

Qy	1	TTGGCGGGCGGAAGCGGCCACAACCCGGCGGATTCGAAAGAGTCTTTAGGAACGCGGTACCA	60
Db	1	TTGGCGGGCGGAAGCGGGCCACAACCCGGCGGATTCGAAAGAGTCTTTAGGAACGCGGTACCA	60
Qy	61	GCGCGGTCTCTCAGAGACAGCGGCCCTTGTCCTCTGTGCGGGCGCGCTCAGCCGCTGCC	120
Db	61	GCGCGGTCTCTCAGAGACAGCGGCCCTTGTCCTCTGTGCGGGCGCGCTCAGCCGCTGCC	120
Qy	121	TCCGCCCTCAGGTTCCTTTTCTAAATCCAAATAAACTTGCACAGGAGCTATGAAAGATT	180
Db	121	TCCGCCCTCAGGTTCCTTTTCTAAATCCAAATAAACTTGCACAGGAGCTATGAAAGATT	180
Qy	181	ATGATGAACCTTCCAATATATATGAATATACATGAACAATATGGGACAGGTGGCTTTGC	240
Db	181	ATGATGAACCTTCCAATATATATGAATATACATGAACAATATGGGACAGGTGGCTTTGC	240
Qy	241	AGGTCAAACTTCGCTGCCATATCCTTTACTCGAGAGATGGTAGCTATATAAAATCATCG	300
Db	241	AGGTCAAACTTCGCTGCCATATCCTTTACTCGAGAGATGGTAGCTATATAAAATCATCG	300
Qy	301	AAAACACACTAGGGAGTGATTTGGCCCGGATCAAAAACGGAGATGGAGCCTTCAGAA	360
Db	301	AAAACACACTAGGGAGTGATTTGGCCCGGATCAAAAACGGAGATGGAGCCTTCAGAA	360
Qy	361	TGAGACATCAGCATATATGTCACATCTACCATGTGCTAGAGACAGCCACAATAATTC	420
Db	361	TGAGACATCAGCATATATGTCACATCTACCATGTGCTAGAGACAGCCACAATAATTC	420
Qy	421	TGGTTCTTGAGTACTGCCCTGGAGGAGAGCTGTTTGACTATATAATTTCCAGGATCG	480
Db	421	TGGTTCTTGAGTACTGCCCTGGAGGAGAGCTGTTTGACTATATAATTTCCAGGATCG	480
Qy	481	TGTCAGAGAGAGAGACCGGGGTGTCTTCCTCGTCAGATAGTATCTGCTGTGCTTATG	540
Db	481	TGTCAGAGAGAGAGACCGGGGTGTCTTCCTCGTCAGATAGTATCTGCTGTGCTTATG	540
Qy	541	ACAGCCAGGGCTATGCTCACAGGACCTCAGCCAGAGAAATTTGCTGTTTGATGAATTC	600
Db	541	ACAGCCAGGGCTATGCTCACAGGACCTCAGCCAGAGAAATTTGCTGTTTGATGAATTC	600
Qy	601	ATAAAATTAAGCTGATTCGACTTTGGTCTCTGTGCMAAACCAAGGGTAAACAAGGATTC	660

Db 1681 GCTCAGTGGAAATTGGATCTCAACCAAGCACATATGGAGGAGACTCCAAAAGAAAGGGAG 1740  
Qy 1741 CCAAAGTGTGGAGAGCTTTGAAGAGGGGTGGATAAGGTTATCACTGTGCTCACCAGGA 1800  
Db 1741 CCAAAGTGTGGAGAGCTTTGAAGAGGGGTGGATAAGGTTATCACTGTGCTCACCAGGA 1800  
Qy 1801 GCAAAAGGAGGGTCTTCGCAGAGACGGGCCCAAGAGACTTAAAGCTTCACTATAATGTGA 1860  
Db 1801 GCAAAAGGAGGGTCTTCGCAGAGACGGGCCCAAGAGACTTAAAGCTTCACTATAATGTGA 1860  
Qy 1861 CTACAACTAGATTAGTGAATCCAGATCAACTGTGTGAATGAATAATGTCTATCTTCCAA 1920  
Db 1861 CTACAACTAGATTAGTGAATCCAGATCAACTGTGTGAATGAATAATGTCTATCTTCCAA 1920  
Qy 1921 AGAAGCATGTTGACTTTGTACAAAAGGGTTATACACTGAAGTGTCAACACACAGTCAAGTT 1980  
Db 1921 AGAAGCATGTTGACTTTGTACAAAAGGGTTATACACTGAAGTGTCAACACACAGTCAAGTT 1980  
Qy 1981 TTGGGAAAGTGACAATGCAATTTGAATTTAGAAGTGTGCCAGCTTCAAAAACCCGATGTGG 2040  
Db 1981 TTGGGAAAGTGACAATGCAATTTGAATTTAGAAGTGTGCCAGCTTCAAAAACCCGATGTGG 2040  
Qy 2041 TGGGTATCAGGAGCGAGCGGCTTAAAGGCGATGCTGGGTTTACAAAAGATTAGTGGAG 2100  
Db 2041 TGGGTATCAGGAGCGAGCGGCTTAAAGGCGATGCTGGGTTTACAAAAGATTAGTGGAG 2100  
Qy 2101 ACATCCTATCTAGCTGCAAGGTATAATTGATGATGATCTTCCATCCTGCCGGATGAGTGG 2160  
Db 2101 ACATCCTATCTAGCTGCAAGGTATAATTGATGATGATCTTCCATCCTGCCGGATGAGTGG 2160  
Qy 2161 GGTGTGATACAGCTCATAAAGACTGTTATGATCGCTTTGATTTTAAAGTTCAITGGAA 2220  
Db 2161 GGTGTGATACAGCTCATAAAGACTGTTATGATCGCTTTGATTTTAAAGTTCAITGGAA 2220  
Qy 2221 CTACCAACTGTTTCTAAAGAGCTATCTTAAAGCAATATCTCTTTGTTTTTAAACAAA 2280  
Db 2221 CTACCAACTGTTTCTAAAGAGCTATCTTAAAGCAATATCTCTTTGTTTTTAAACAAA 2280  
Qy 2281 GATATTATTTGTGATGAATCAATCAAGCCCATCTGCATATGTTACTGTCTTTT 2340  
Db 2281 GATATTATTTGTGATGAATCAATCAAGCCCATCTGCATATGTTACTGTCTTTT 2340  
Qy 2341 TAATCATGTGTTTGTATATTAATAATTGTTGACTTTCTTAGATTCACTTCCATATGTG 2400  
Db 2341 TAATCATGTGTTTGTATATTAATAATTGTTGACTTTCTTAGATTCACTTCCATATGTG 2400  
Qy 2401 AATGTAAGCTCTTAACATATGCTCTTTGTAATGTGTAATTTCTTCTGAAATTAACCAT 2460  
Db 2401 AATGTAAGCTCTTAACATATGCTCTTTGTAATGTGTAATTTCTTCTGAAATTAACCAT 2460  
Qy 2461 TTGTGAATAT 2470  
Db 2461 TTGTGAATAT 2470

RESULT 5  
ADB80562  
ID ADB80562 standard; DNA; 2470 BP.  
XX  
AC ADB80562;  
XX  
DT 04-DEC-2003 (first entry)  
XX  
DE Ovarian cancer-associated transcript #64.  
XX  
KW cytostatic; gene therapy; vaccine; ovarian cancer; diagnosis;  
KW post-operative chemotherapy; radiation therapy; tumour prognosis;  
KW pre-cancerous lesion detection; ds; gene.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT CDS 171..2126

FT XX /\*tag= a  
PN W02002102235-A2.  
XX  
PD 27-DEC-2002.  
XX  
PF 18-JUN-2002; 2002WO-US019297.  
XX  
PR 18-JUN-2001; 2001US-0299234P.  
PR 27-AUG-2001; 2001US-0315287P.  
PR 05-SEP-2001; 2001US-0317544P.  
PR 13-NOV-2001; 2001US-0350666P.  
PR 12-APR-2002; 2002US-0372246P.  
XX  
PA (EOSB-) EOS BIOTECHNOLOGY INC.  
XX  
PI Mack DH, Gish KC;  
XX  
DR WPI; 2003-167431/16.  
XX  
PT P-PSDB; ADB80563.  
XX  
PT Detecting an ovarian cancer-associated transcript in a cell from a  
patient, comprises contacting a biological sample from the patient with a  
polynucleotide that hybridizes to an ovarian cancer gene.  
XX  
PS Claim 10; Page 318; 332pp; English.  
XX  
CC The invention relates to a method of detecting an ovarian cancer-  
associated transcript in a cell from a patient, by contacting a  
biological sample from the patient with a polynucleotide that selectively  
hybridizes to a sequence at least 80% identical to any of one of 80  
nucleic acid sequences given in the specification. The method is useful  
in diagnosing ovarian cancer and in identifying and using agents and/or  
targets that inhibit ovarian cancer. The nucleic acid molecule,  
polypeptide and the antibody may also be used in detecting ovarian  
cancers, monitoring and early detection of relapse following treatment,  
chemotherapy or radiation therapy, in selection of pre-cancerous lesions,  
and as vaccines. This sequence corresponds to one of the nucleic acids  
used for the detection method of the invention.  
XX  
SQ Sequence 2470 BP; 767 A; 488 C; 536 G; 679 T; 0 U; 0 Other;  
Query Match 100.0%; Score 2470; DB 10; Length 2470;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2470; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 TTGGCGGGCGGAAGCGGCCACAAACCCGGCGATCGAAAAGATTCTTAGGAACGCCGTACCA 60  
Db 1 TTGGCGGGCGGAAGCGGCCACAAACCCGGCGATCGAAAAGATTCTTAGGAACGCCGTACCA 60  
Qy 61 GCGCGTCTCTCAGGACAGCAGCGCCCTGCTCTCTGTCGGCGCGCTCAGCCGTGCC 120  
Db 61 GCGCGTCTCTCAGGACAGCAGCGCCCTGCTCTCTGTCGGCGCGCTCAGCCGTGCC 120  
Qy 121 TCGCCCTCCTCAGGTCTTTTCTTAATTCAAAATAAATTCGAAGAGACTATGAAGAATT 180  
Db 121 TCGCCCTCCTCAGGTCTTTTCTTAATTCAAAATAAATTCGAAGAGACTATGAAGAATT 180  
Qy 181 ATGATGAACCTTCTCAAAATATTATGAATTACATGAATACTATTGGACAGGTGGCTTTGCAA 240  
Db 181 ATGATGAACCTTCTCAAAATATTATGAATTACATGAATACTATTGGACAGGTGGCTTTGCAA 240  
Qy 241 AGGTCAAACTTGCCTGCCATATCCTTACTGGAGAGATGGTAGCTATAAAAATCATGGATA 300  
Db 241 AGGTCAAACTTGCCTGCCATATCCTTACTGGAGAGATGGTAGCTATAAAAATCATGGATA 300  
Qy 301 AAAACACACTAGGAGTGAATTTGCCCGGATCAAAACCGAGATTGAGGCTTTGAAGAACC 360  
Db 301 AAAACACACTAGGAGTGAATTTGCCCGGATCAAAACCGAGATTGAGGCTTTGAAGAACC 360  
Qy 361 TGAGACATCAGCATATATGTCAACTCTACCTGTGTAGAGACAGCAACAAATATTCA 420

Db	361	TGAGCATCAGCATATATATGTCAACTCTACATGTGCTAGAGACGCAACAAAATATTC	420
Qy	421	TGTTCTTGTAGTACTGCCCTCGAGGAGAGCTGTTTGACTATATAATTTCCAGGATCGCC	480
Db	421	TGTTCTTGTAGTACTGCCCTCGAGGAGAGCTGTTTGACTATATAATTTCCAGGATCGCC	480
Qy	481	TGTCAGAAAGAGAGACCGGGTGTCTTCGTCAGATAGTATCTGCTGTGCTTATGTGC	540
Db	481	TGTCAGAAAGAGAGACCGGGTGTCTTCGTCAGATAGTATCTGCTGTGCTTATGTGC	540
Qy	541	ACAGCCAGGCTATGCTCAAGGAGCTCAAGCCAGCAAAATTTGCTGTTTGAATATATC	600
Db	541	ACAGCCAGGCTATGCTCAAGGAGCTCAAGCCAGCAAAATTTGCTGTTTGAATATATC	600
Qy	601	ATAAATTAAGCTGATTTGACTTTTGGTCTCTGTGCAAAACCCAGGGTAAACAAGGATTAAC	660
Db	601	ATAAATTAAGCTGATTTGACTTTTGGTCTCTGTGCAAAACCCAGGGTAAACAAGGATTAAC	660
Qy	661	ATCTACAGATGCTGTGGAGTCTGGCTTATGTCAGCACCTGATGTTAATACAGGCAAT	720
Db	661	ATCTACAGATGCTGTGGAGTCTGGCTTATGTCAGCACCTGATGTTAATACAGGCAAT	720
Qy	721	CATATCTTGATCAGAGGAGATGTTTGGAGCATGGGCATCTGTTATATGTTCTTATGT	780
Db	721	CATATCTTGATCAGAGGAGATGTTTGGAGCATGGGCATCTGTTATATGTTCTTATGT	780
Qy	781	GTGGATTTCTACCATTTGATGATGATTAATGTAATGGCTTTATACAGAAAGATTAAGAG	840
Db	781	GTGGATTTCTACCATTTGATGATGATTAATGTAATGGCTTTATACAGAAAGATTAAGAG	840
Qy	841	GAAATATGATGTTCCCAAGTGCTCTCCAGATGATGATTCCTGCTTCTTCAACAAATGC	900
Db	841	GAAATATGATGTTCCCAAGTGCTCTCCAGATGATGATTCCTGCTTCTTCAACAAATGC	900
Qy	901	TGAGGTGACCAAGAAACGGATTTCTATGAAATCTATTTGAACCATCCCTCGATCA	960
Db	901	TGAGGTGACCAAGAAACGGATTTCTATGAAATCTATTTGAACCATCCCTCGATCA	960
Qy	961	TGCAAGATTACAACTATCTGTTGAGTGCAAGCAAGAAATCTTTTATTCACCTCGATG	1020
Db	961	TGCAAGATTACAACTATCTGTTGAGTGCAAGCAAGAAATCTTTTATTCACCTCGATG	1020
Qy	1021	ATGATTGCGTAACAGAACTTTCTGTATACATCAAGAAACAAAGGCAAAATGAGGAT	1080
Db	1021	ATGATTGCGTAACAGAACTTTCTGTATACATCAAGAAACAAAGGCAAAATGAGGAT	1080
Qy	1081	TAAATTCAGTGGCAGTATGATCACCTCACGGCTACCTATCTTCTGCTTCTAGCCAAG	1140
Db	1081	TAAATTCAGTGGCAGTATGATCACCTCACGGCTACCTATCTTCTGCTTCTAGCCAAG	1140
Qy	1141	AGGCTCGGGGAAACCAAGTTCTGTTTAAAGCTTTCTTTCTTCTGTCGACCAAGCAGTG	1200
Db	1141	AGGCTCGGGGAAACCAAGTTCTGTTTAAAGCTTTCTTTCTTCTGTCGACCAAGCAGTG	1200
Qy	1201	CTACCCCAATTCACAGATCAAGTCAAAATTAATGAGTCTGGAAGATGTGACCGCAAGTG	1260
Db	1201	CTACCCCAATTCACAGATCAAGTCAAAATTAATGAGTCTGGAAGATGTGACCGCAAGTG	1260
Qy	1261	ATAAATAATTAATGCGGGATTAATAGACTATGATTTGGTGAAGATGATTTATCAACAG	1320
Db	1261	ATAAATAATTAATGCGGGATTAATAGACTATGATTTGGTGAAGATGATTTATCAACAG	1320
Qy	1321	GTGCTGCTACTCCCGAAACATCACAGTTTACCAAGTACTGGAAGATCAAAATGGGGTGG	1380
Db	1321	GTGCTGCTACTCCCGAAACATCACAGTTTACCAAGTACTGGAAGATCAAAATGGGGTGG	1380
Qy	1381	AATCTAAATCATTAATCCAGCTTATGAGAAACACCTGCAAAATAAATTAAGAAACAAG	1440
Db	1381	AATCTAAATCATTAATCCAGCTTATGAGAAACACCTGCAAAATAAATTAAGAAACAAG	1440
Qy	1441	AAATGTATATCTCTTAAGTCTGCTGTAAGAAATGAAGATGATCTTTATGTTCTTGAGC	1500

RESULT 6  
ADE38346

Db	1441	AAATGTATATCTCTTAAGTCTGCTGTAAGAAATGAAGATGATCTTTATGTTCTTGAGC	1500
Qy	1501	CAAAGACTCCAGTTAATTAAGAACCAAGCATAAAGAGAAATATCTCACTAGCCCAATCGTT	1560
Db	1501	CAAAGACTCCAGTTAATTAAGAACCAAGCATAAAGAGAAATATCTCACTAGCCCAATCGTT	1560
Qy	1561	ACACTACACCTTCAAAAGCTAGAAAACCAAGTCTGCTGAAAGAAATCTCAATTAATAACAG	1620
Db	1561	ACACTACACCTTCAAAAGCTAGAAAACCAAGTCTGCTGAAAGAAATCTCAATTAATAACAG	1620
Qy	1621	TAAATTTCAACAGGAACAGACAAGTTAATGACAGGTGTCTATTAGCCCTGAGAGCCGTGCC	1680
Db	1621	TAAATTTCAACAGGAACAGACAAGTTAATGACAGGTGTCTATTAGCCCTGAGAGCCGTGCC	1680
Qy	1681	GCTCAGTGGAAATGGATCTCAACCAAGCACATATGAGAGAGACTCCAAAAGAAAGGGAG	1740
Db	1681	GCTCAGTGGAAATGGATCTCAACCAAGCACATATGAGAGAGACTCCAAAAGAAAGGGAG	1740
Qy	1741	CCAAAGTGTGGAGCCCTTGAAGGGGTGGATAAGGTTATCACTGTGCTCAACAGGA	1800
Db	1741	CCAAAGTGTGGAGCCCTTGAAGGGGTGGATAAGGTTATCACTGTGCTCAACAGGA	1800
Qy	1801	GCAAAAGGAAGGTTCTGCGAGAGACGGGCCGAGAACTTAAAGCTTCACTATAATGTA	1860
Db	1801	GCAAAAGGAAGGTTCTGCGAGAGACGGGCCGAGAACTTAAAGCTTCACTATAATGTA	1860
Qy	1861	CTACAACTAGATGATGAAATCCAGATCAACTGTTGAAATGAATTAATGTCTATTTCCAA	1920
Db	1861	CTACAACTAGATGATGAAATCCAGATCAACTGTTGAAATGAATTAATGTCTATTTCCAA	1920
Qy	1921	AGAGCATGTTGACTTTGTACAAAAGGGTTATACACTGAAAGTGTCAACACAGTCAGATT	1980
Db	1921	AGAGCATGTTGACTTTGTACAAAAGGGTTATACACTGAAAGTGTCAACACAGTCAGATT	1980
Qy	1981	TTGGGAAAGTGCACATGCAATTTTGAATTTAGAAGTGTGCGAGCTTCAAAAACCCGATGG	2040
Db	1981	TTGGGAAAGTGCACATGCAATTTTGAATTTAGAAGTGTGCGAGCTTCAAAAACCCGATGG	2040
Qy	2041	TGGGTATCAGGAGGCGGCTTAAAGGCGATGCTGGGTTTACAAAAGATTAAGTGAAG	2100
Db	2041	TGGGTATCAGGAGGCGGCTTAAAGGCGATGCTGGGTTTACAAAAGATTAAGTGAAG	2100
Qy	2101	ACATCTATCTAGTGCAGGATTAATTTGATGAGATTTCTTCCATCTGCGGATGAGTGTG	2160
Db	2101	ACATCTATCTAGTGCAGGATTAATTTGATGAGATTTCTTCCATCTGCGGATGAGTGTG	2160
Qy	2161	GGTGTGATACAGCTTACATAAAGTGTATGATGCTTTGATTTTAAAGTTTCAATGGAA	2220
Db	2161	GGTGTGATACAGCTTACATAAAGTGTATGATGCTTTGATTTTAAAGTTTCAATGGAA	2220
Qy	2221	CTACCAACTTGTCTTAAAGAGCTATCTTAAAGCAATATCTCTTTGTTTAAACAAA	2280
Db	2221	CTACCAACTTGTCTTAAAGAGCTATCTTAAAGCAATATCTCTTTGTTTAAACAAA	2280
Qy	2281	GATATATTTTGTGATGAATCTAAATCAAGCCCATCTGTCATTAATGTTACTGTCTTTT	2340
Db	2281	GATATATTTTGTGATGAATCTAAATCAAGCCCATCTGTCATTAATGTTACTGTCTTTT	2340
Qy	2341	TAAATCATGCTGTTTGTATATTAATTAATTTGATGCTTTCTTAGATTTCACTTCCATATG	2400
Db	2341	TAAATCATGCTGTTTGTATATTAATTAATTTGATGCTTTCTTAGATTTCACTTCCATATG	2400
Qy	2401	AATGTAAGCTTAACTATGCTCTTTGTAATGTAATTTCTTTCTGAAATAAACCAT	2460
Db	2401	AATGTAAGCTTAACTATGCTCTTTGTAATGTAATTTCTTTCTGAAATAAACCAT	2460
Qy	2461	TTGTGAATAT	2470
Db	2461	TTGTGAATAT	2470

ID ADE38346 standard; DNA; 2470 BP.  
AC ADE38346;  
XX  
XX  
XX 29-JAN-2004 (first entry)  
XX  
XX Human protein 2089 gene sequence.  
DE  
DE  
XX tumorigenic disorder; angiogenic disorder; aberrant gene expression;  
KW aberrant protein activity; cytostatic; antithyroid; antidiabetic;  
KW ophthalmological; cancer; breast cancer; colon cancer; lung cancer;  
KW prostatic cancer; Grave's disease; diabetic retinopathy; gene; ds;  
KW protein 2089.  
XX  
XX Homo sapiens.  
OS  
XX  
XX Key Location/Qualifiers  
FH 171.2126  
FT CDS /\*tag= a  
FT FT /product= "Human protein 2089"  
FT  
XX  
XX WO2003065006-A2.  
XX  
XX 07-AUG-2003.  
XX  
XX 30-JAN-2003; 2003WO-US002588.  
XX  
XX 31-JAN-2002; 2002US-0353600P.  
XX 15-MAR-2002; 2002US-0364517P.  
XX 09-APR-2002; 2002US-0371075P.  
XX 10-APR-2002; 2002US-0371507P.  
XX 16-APR-2002; 2002US-0372984P.  
XX 19-APR-2002; 2002US-0374194P.  
XX 24-MAY-2002; 2002US-0382955P.  
XX 31-MAY-2002; 2002US-0385023P.  
XX 14-JUN-2002; 2002US-0388533P.  
XX 17-JUN-2002; 2002US-0389395P.  
XX 25-JUN-2002; 2002US-0391324P.  
XX 15-JUL-2002; 2002US-0395944P.  
XX 22-JUL-2002; 2002US-0397726P.  
XX 13-AUG-2002; 2002US-0403046P.  
XX 22-AUG-2002; 2002US-0405155P.  
XX 27-AUG-2002; 2002US-0406361P.  
XX 25-OCT-2002; 2002US-0421195P.  
XX 12-NOV-2002; 2002US-0425456P.  
XX 19-NOV-2002; 2002US-0427626P.  
XX 10-DEC-2002; 2002US-0432122P.  
XX  
XX (MILL-) MILLENNIUM PHARM INC.  
XX  
XX Hunter JJ, Macbeth KJ, Tsai P, Lesoon A, Lightcap ES;  
PI Williamson MW, Rudolph-Owen LA;  
PI  
XX  
XX WPI; 2003-646176/61.  
XX P-PSDB; ADE38347.  
XX  
XX Treating subject having tumorigenic disorder or angiogenic disorder  
PT caused by aberrant polypeptide e.g., N-formylpeptide receptor or nucleic  
PT acid, by administering a modulator.  
XX  
XX Disclosure; SEQ ID NO 7; 454pp; English.  
XX  
XX This invention relates to a novel method of treating a human subject  
CC having a tumorigenic disorder or angiogenic disorder, caused by aberrant  
CC gene expression or activity of an isolated protein, by administering a  
CC modulator. The modulator may have cytostatic, antithyroid, antidiabetic  
CC or ophthalmological activity. The method is useful for treating a subject  
CC having a tumorigenic or angiogenic disorder, in particular for treating  
CC cancer (for example breast cancer, colon cancer, lung cancer or prostatic  
CC cancer) and, for example, Grave's disease and diabetic retinopathy. The  
CC present sequence is a DNA sequence which encodes the novel isolated human  
CC protein 2089 of the invention.  
XX

SQ Sequence 2470 BP; 767 A; 488 C; 536 G; 679 T; 0 U; 0 Other;  
Query Match 100.0%; Score 2470; DB 10; Length 2470;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2470; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
XX  
QY 1 TTGGCGGGCGGAAGCGGCCACAAACCCGGCGATCGAAAAAGATTCTTTAGGAACCCCGTACCA 60  
DB 1 TTGGCGGGCGGAAGCGGCCACAAACCCGGCGATCGAAAAAGATTCTTTAGGAACCCCGTACCA 60  
QY 61 GCCGGTCTCTCAGACAGCAGGCCCTCTGCTCTTCTGTGGCGCGCTCAGCCGTCGCC 120  
DB 61 GCCGGTCTCTCAGACAGCAGGCCCTCTGCTCTTCTGTGGCGCGCTCAGCCGTCGCC 120  
QY 121 TCCGCCCTCAGGTTCTTTTCTTAATCCAAATAAATTTGCAAGAGGACTATGAAGATT 180  
DB 121 TCCGCCCTCAGGTTCTTTTCTTAATCCAAATAAATTTGCAAGAGGACTATGAAGATT 180  
QY 181 ATGATGAACCTTCTCAAAATATTATGAATTAACATGAATAATTGGGACAGGTGGCTTGC 240  
DB 181 ATGATGAACCTTCTCAAAATATTATGAATTAACATGAATAATTGGGACAGGTGGCTTGC 240  
QY 241 AGGTCAAACTTGCCTGCCATATCTTCTAGTGGAGAGATGGTACCTATAAATCATGGATA 300  
DB 241 AGGTCAAACTTGCCTGCCATATCTTCTAGTGGAGAGATGGTACCTATAAATCATGGATA 300  
QY 301 AAAACACACTAGGGAGTGATTTGCCCGCGATCAAAACCGAGATTGAGGCTTTGAAGAACC 360  
DB 301 AAAACACACTAGGGAGTGATTTGCCCGCGATCAAAACCGAGATTGAGGCTTTGAAGAACC 360  
QY 361 TGAGACATCAGCATATATGTCAACTCTACCATGTCTAGAGACAGCCAAATAATTCA 420  
DB 361 TGAGACATCAGCATATATGTCAACTCTACCATGTCTAGAGACAGCCAAATAATTCA 420  
QY 421 TGGTCTTTGAGTACTGCCCTGGAGGAGAGCTGTTTGAATTAATTTCCAGGATCGCC 480  
DB 421 TGGTCTTTGAGTACTGCCCTGGAGGAGAGCTGTTTGAATTAATTTCCAGGATCGCC 480  
QY 481 TGTCTAGAGAGGAGACCCGGGTTGCTTCCGTCAGATAGTATCTGCTGTGTGTTATGTC 540  
DB 481 TGTCTAGAGAGGAGACCCGGGTTGCTTCCGTCAGATAGTATCTGCTGTGTGTTATGTC 540  
QY 541 ACAGCCAGGGCTATGCTCAAGGACCTCAAGCCAGAAAAATTTGCTGTTGATGAATATC 600  
DB 541 ACAGCCAGGGCTATGCTCAAGGACCTCAAGCCAGAAAAATTTGCTGTTGATGAATATC 600  
QY 601 ATAAATTAAGCTGATGATCTTTGGTCTCTGTGCAAAACCCAGGGTAACAGGATTACC 660  
DB 601 ATAAATTAAGCTGATGATCTTTGGTCTCTGTGCAAAACCCAGGGTAACAGGATTACC 660  
QY 661 ATCTACAGACATGCTGTGGAGTCTGGCTTATGTCAGACACCTGAGTTAATACAAAGCAAT 720  
DB 661 ATCTACAGACATGCTGTGGAGTCTGGCTTATGTCAGACACCTGAGTTAATACAAAGCAAT 720  
QY 721 CATATCTTGGATCAGAGCAGATGTTTGGAGCATGGGCATCTGTTATATGTTCTTATGT 780  
DB 721 CATATCTTGGATCAGAGCAGATGTTTGGAGCATGGGCATCTGTTATATGTTCTTATGT 780  
QY 781 GTGGATTTCTACCATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 840  
DB 781 GTGGATTTCTACCATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 840  
QY 841 GAAAAATATCATGTTCCCAAGTGGCTCTCTCCAGTAGCATCTCTGCTTCTTCAACAAATGC 900  
DB 841 GAAAAATATCATGTTCCCAAGTGGCTCTCTCCAGTAGCATCTCTGCTTCTTCAACAAATGC 900  
QY 901 TGCAGGTGGACCCAAAGAAACCGGATTTCTATGAAAAATCTATTGAAACCATCCCTGATCA 960  
DB 901 TGCAGGTGGACCCAAAGAAACCGGATTTCTATGAAAAATCTATTGAAACCATCCCTGATCA 960  
QY 961 TGCAGATTTACAACTATCTCTGTTGAGTGGCAAGCAAGATCCTTTTATTCACCTCGATG 1020  
DB 961 TGCAGATTTACAACTATCTCTGTTGAGTGGCAAGCAAGATCCTTTTATTCACCTCGATG 1020

QY 1021 ATGATTGGCTAAGAGAACTTTCTGTATCATCAGAGAAACAGGCAAAACAATGGAGATT 1080  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
QY 1021 ATGATTGGCTAAGAGAACTTTCTGTATCATCAGAGAAACAGGCAAAACAATGGAGATT 1080  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
QY 1081 TAAATTTCACTGTGGCAGTAGATGACCTCAGCGGTACCTATCTTCTGCTTCTAGCCAGA 1140  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
QY 1081 TAAATTTCACTGTGGCAGTAGATGACCTCAGCGGTACCTATCTTCTGCTTCTAGCCAGA 1140  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
QY 1141 AGGCTCGGGGAAAACAGTTGCTTTAAGGCTTTCTTTCTCTCTGTGACAGCCAGTG 1200  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
QY 1201 CTACCCCATTCACAGACATCAAGTCAAAATTTGGAGTCTGGAGATGTGACCGCAAGTG 1260  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
QY 1201 CTACCCCATTCACAGACATCAAGTCAAAATTTGGAGTCTGGAGATGTGACCGCAAGTG 1260  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
QY 1261 ATAAAAATTTATGTGGCGGATTAATAGACTATGATTTGGTGTGAAGATGATTTATCAACAG 1320  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
QY 1261 ATAAAAATTTATGTGGCGGATTAATAGACTATGATTTGGTGTGAAGATGATTTATCAACAG 1320  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
QY 1321 GTGCTGCTACTCCCGAATCATCAGTTTACCAAGTACTGGACAGATCAAAATGGGGTGG 1380  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
QY 1321 GTGCTGCTACTCCCGAATCATCAGTTTACCAAGTACTGGACAGATCAAAATGGGGTGG 1380  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
QY 1381 AATCTAAATCATTAATCTCAGCTTATGACAGAACCTGCAAAATTAATTAAGAACAAAG 1440  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
QY 1381 AATCTAAATCATTAATCTCAGCTTATGACAGAACCTGCAAAATTAATTAAGAACAAAG 1440  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
QY 1441 AAAATGTATATCTCTCTAAGTCTGCTGTAAGAAAGTAAAGTAAAGTAAAGTAAAGTAAAG 1500  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
QY 1441 AAAATGTATATCTCTCTAAGTCTGCTGTAAGAAAGTAAAGTAAAGTAAAGTAAAGTAAAG 1500  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
QY 1501 CAAAGACTCCAGTTAATAGAAACCAAGCATAAGAGAGAAATCTCACTACGCCAAATCGTT 1560  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
QY 1501 CAAAGACTCCAGTTAATAGAAACCAAGCATAAGAGAGAAATCTCACTACGCCAAATCGTT 1560  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
QY 1561 ACATACACCTTCAAAAGCTAGAAACCAAGTCTGCTGAAAGAAATCTCAATTAATAATACCAAG 1620  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
QY 1561 ACATACACCTTCAAAAGCTAGAAACCAAGTCTGCTGAAAGAAATCTCAATTAATAATACCAAG 1620  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
QY 1621 TAAATTTCAACAGGAAACAGACAGTTAATCAGAGGTGTCAATGACCTGAGAGCGGTGCC 1680  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
QY 1621 TAAATTTCAACAGGAAACAGACAGTTAATCAGAGGTGTCAATGACCTGAGAGCGGTGCC 1680  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
QY 1681 GCTCAGTGAATTTGGATCTCAACCAAGCACAATATGAGAGAGACTCCAAAAGAAAGGGAG 1740  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
QY 1681 GCTCAGTGAATTTGGATCTCAACCAAGCACAATATGAGAGAGACTCCAAAAGAAAGGGAG 1740  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
QY 1741 CCAAGAGTTTGGAGCCTTGAAGAGGGGTTGGATAAGTTATCACTGTGCTCACCAAGGA 1800  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
QY 1741 CCAAGAGTTTGGAGCCTTGAAGAGGGGTTGGATAAGTTATCACTGTGCTCACCAAGGA 1800  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
QY 1801 GCAAGAGGAGGTTTCTGACAGAGCGGCCAGAGACTCAAGCTTCACTATATATGGA 1860  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
QY 1801 GCAAGAGGAGGTTTCTGACAGAGCGGCCAGAGACTCAAGCTTCACTATATATGGA 1860  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
QY 1861 CTACAACTAGATTAGTGAATCCAGATCACTGTGTAATGAATATGCTATCTTCCAA 1920  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
QY 1861 CTACAACTAGATTAGTGAATCCAGATCACTGTGTAATGAATATGCTATCTTCCAA 1920  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
QY 1921 AGAAGCATGTTGATTTGTAACAAAAGGGTTATACACTGAAGTGTCAACACAGTCAAGTT 1980  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
QY 1921 AGAAGCATGTTGATTTGTAACAAAAGGGTTATACACTGAAGTGTCAACACAGTCAAGTT 1980  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
QY 1981 TTGGGAAAGTGAATGCAATTTGTAATTAAGAGTGTCCAGCTTCAAAAACCCGATGGG 2040  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
QY 1981 TTGGGAAAGTGAATGCAATTTGTAATTAAGAGTGTCCAGCTTCAAAAACCCGATGGG 2040  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
QY 2041 TGGGTATCAGGAGGCGGCTTAAAGGCGATGCTCGGTTTACAAAAGATTAGTGGAG 2100  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
QY 2041 TGGGTATCAGGAGGCGGCTTAAAGGCGATGCTCGGTTTACAAAAGATTAGTGGAG 2100  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

QY 2101 ACATCTATCTAGCTGCAAGGTATTAATTCATGATGATCTTCCATCTCCGCGATGAGTGTG 2160  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
QY 2101 ACATCTATCTAGCTGCAAGGTATTAATTCATGATGATCTTCCATCTCCGCGATGAGTGTG 2160  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
QY 2161 GGTGTGATACAGCTTACATAAAGACTGTGATGATCGCTTTGATTTTAAAGTTTCATTGGAA 2220  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
QY 2161 GGTGTGATACAGCTTACATAAAGACTGTGATGATCGCTTTGATTTTAAAGTTTCATTGGAA 2220  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
QY 2221 CTACCAACTTGTCTTAAAGAGCTATCTTTAAGACCAATATCTCTTTGTTTTTAAACAAA 2280  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
QY 2221 CTACCAACTTGTCTTAAAGAGCTATCTTTAAGACCAATATCTCTTTGTTTTTAAACAAA 2280  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
QY 2281 GATATATTTTGTGATGATGATCTTAATCAAGCCCATCTGTCATTATGTTACTGTCTTTT 2340  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
QY 2281 GATATATTTTGTGATGATGATCTTAATCAAGCCCATCTGTCATTATGTTACTGTCTTTT 2340  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
QY 2341 TAATCATGTGGTTTTGTATATTAATAATTTGTTGACTTTCTTAGATTTCACCTCCATATGTG 2400  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
QY 2341 TAATCATGTGGTTTTGTATATTAATAATTTGTTGACTTTCTTAGATTTCACCTCCATATGTG 2400  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
QY 2401 AATGTAAGCTTTAACTACTGCTCTTTGTAATGTAATTTCTTTCTGAAAATAAACCAT 2460  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
QY 2401 AATGTAAGCTTTAACTACTGCTCTTTGTAATGTAATTTCTTTCTGAAAATAAACCAT 2460  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
QY 2461 TTGTGAATAT 2470  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
QY 2461 TTGTGAATAT 2470  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
RESULT 7  
ADN39187  
ID ADN39187 standard; cDNA; 2470 BP.  
XX  
AC ADN39187;  
XX  
DT 17-JUN-2004 (first entry)  
XX  
DE Cancer/angiogenesis/fibrosis-related nucleic acid, SEQ ID NO:505.  
XX  
KW Human; differential expression; cancer; angiogenic disorder;  
KW fibrotic disorder; psoriasis; ischaemia; heart disease; atherosclerosis;  
KW inflammatory disease; autoimmune disease;  
KW retinal neovascularization syndrome; scarring; uterine fibroid;  
KW detection; diagnosis; prognosis; drug screening; drug targeting;  
KW wound healing; contraception; cytostatic; cardiant; immunomodulatory;  
KW vulnery; gene therapy; vaccine; gene; ss.  
XX  
OS Homo sapiens.  
XX  
PN WO2003042661-A2.  
XX  
PD 22-MAY-2003.  
XX  
PF 13-NOV-2002; 2002WO-US036810.  
XX  
PR 13-NOV-2001; 2001US-0350666P.  
PR 21-NOV-2001; 2001US-0332464P.  
PR 29-NOV-2001; 2001US-0334393P.  
PR 03-DEC-2001; 2001US-0335394P.  
PR 14-DEC-2001; 2001US-0340376P.  
PR 08-JAN-2002; 2002US-0347211P.  
PR 10-JAN-2002; 2002US-0347349P.  
PR 08-FEB-2002; 2002US-035250P.  
PR 13-FEB-2002; 2002US-0356714P.  
PR 20-FEB-2002; 2002US-0359077P.  
PR 29-MAR-2002; 2002US-0368809P.  
PR 04-APR-2002; 2002US-0370110P.  
PR 12-APR-2002; 2002US-0372246P.  
PR 05-JUN-2002; 2002US-0386614P.  
PR 16-JUL-2002; 2002US-0396839P.  
PR 22-JUL-2002; 2002US-039775P.  
PR 22-JUL-2002; 2002US-0397845P.  
PR 09-SEP-2002; 2002US-0409450P.



XX	(BOSB-) EOS BIOTECHNOLOGY INC.
PA	Afar D, Aziz N, Ginsburg WM, Gish KC, Glynn R, Hevezi PA;
XX	Mack DH, Murray R, Watson SR, Wilson KE, Ziolknik A;
PI	
PI	
XX	WPI; 2003-468649/44.
DR	P-PSDB; ADN39188.
DR	
XX	Determining the presence or absence of a pathological cell in a patient,
PT	useful for diagnosing, prognosing or treating cancer, comprises detecting
PT	a nucleic acid in a biological sample.
PT	
XX	Claim 8; SEQ ID NO 505; 1385pp; English.
PS	
XX	The invention relates to nucleic acids and proteins (ADN38683-ADN40064)
CC	whose expression is upregulated or downregulated in specific cancers or
CC	other diseases such as angiogenic or fibrotic disorders, and to methods
CC	of determining the presence or absence of a pathological cell in a
CC	patient by detecting a nucleic acid at least 80% identical to those of
CC	the invention or by detecting a polypeptide of the invention. The
CC	invention also relates to expression vectors and host cells comprising a
CC	nucleic acid of the invention; antibodies which specifically bind a
CC	polypeptide of the invention; use of such antibodies for drug targeting;
CC	and methods of screening for modulators of activity or expression of the
CC	antibodies and nucleic acids. The nucleic acids, polypeptides,
CC	antibodies and methods are useful for diagnosing, prognosing and treating
CC	cancer and other conditions such as psoriasis, ischaemia, heart disease,
CC	atherosclerosis, inflammatory diseases, autoimmune diseases, retinal
CC	neovascularisation syndromes, scarring and uterine fibroids. They may
CC	also be useful in wound healing and in contraception. The present
CC	sequence represents a nucleic acid sequence of the invention.
XX	
SQ	Sequence 2470 BP; 767 A; 488 C; 536 G; 679 T; 0 U; 0 Other;
	Query Match 100.0%; Score 2470; DB 11; Length 2470;
	Best Local Similarity 100.0%; Pred. No. 0;
	Matches 2470; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	1 TTGGCGGCGGAAGCGGCACAAACCGCGCATCGAAGAAGATTTCTTAGGAACGCCGTACCA 60
DB	1 TTGGCGGCGGAAGCGGCACAAACCGCGCATCGAAGAAGATTTCTTAGGAACGCCGTACCA 60
QY	61 GC CGCGTCTCTCAGGACAGACGCGCCCTGTCTCTCTGTCGGCGCGCGCTCAGCGCGTGCCC 120
DB	61 GC CGCGTCTCTCAGGACAGACGCGCCCTGTCTCTCTGTCGGCGCGCGCTCAGCGCGTGCCC 120
QY	121 TC GCCGCCCTCAGGTTCTTTTTCTTAATTC AAATAACTTTGCAAGAGGACTATGAAGAATT 180
DB	121 TC GCCGCCCTCAGGTTCTTTTTCTTAATTC CAATAAATCTTCCAAGAGGACTATGAAGAATT 180
QY	181 ATGATGAATCTTCAAATATTAATGAATTAATCATGAATTAATTTGGGACAGGTGGCTTGCAA 240
DB	181 ATGATGAATCTTCAAATATTAATGAATTAATCATGAATTAATTTGGGACAGGTGGCTTGCAA 240
QY	241 AGGTCAAACTTCCTGCCATATCTTTACTTGGAGAGATGTAGCTATATAAAAATCATGGATA 300
DB	241 AGGTCAAACTTCCTGCCATATCTTTACTTGGAGAGATGTAGCTATATAAAAATCATGGATA 300
QY	301 AAAACAACACTAGGGAGTGATTTGGCCCCGGGATCAAAACCGGAGATTGAGSCCTTTGAAGAACC 360
DB	301 AAAACAACACTAGGGAGTGATTTGGCCCCGGGATCAAAACCGGAGATTGAGSCCTTTGAAGAACC 360
QY	361 TGAGACATCAGCATATATGTCACATCTTACCATGTGCTAGAGACAGCCACAAATATTTCA 420
DB	361 TGAGACATCAGCATATATGTCACATCTTACCATGTGCTAGAGACAGCCACAAATATTTCA 420
QY	421 TGGTCTCTCAGTACTGCCCTGGAGGAGAGCTGCTTTTGACTATATATATTTCCCGAGATCGCC 480
DB	421 TGGTCTCTCAGTACTGCCCTGGAGGAGAGCTGCTTTTGACTATATATATTTCCCGAGATCGCC 480
QY	481 TGT'CAGAAGAGGAGACCCGGGTTGTCTTCGTCAGATAGTATCTGCTGTCTTATGTGC 540
DB	TGT'CAGAAGAGGAGACCCGGGTTGTCTTCGTCAGATAGTATCTGCTGTCTTATGTGC 540

```
QY 1621 TAAATTCACAGGACAGACAGCTTAATGACAGGTGTCTATTAGCCTTGAGAGCGGTGCC 1680
Db 1621 TAAATTCACAGGACAGACAGCTTAATGACAGGTGTCTATTAGCCTTGAGAGCGGTGCC 1680
QY 1681 GCTCAGTGGAAATTTGGATCTCAACCAAGCACATATGGAGGAGACTCCAAAAGAAAGGGAG 1740
Db 1681 GCTCAGTGGAAATTTGGATCTCAACCAAGCACATATGGAGGAGACTCCAAAAGAAAGGGAG 1740
QY 1741 CCAAAGTGTGGAGGCTTTGAAAGGGGTTGGATAAGGTTATCACTGTGCTCACCAGGA 1800
Db 1741 CCAAAGTGTGGAGGCTTTGAAAGGGGTTGGATAAGGTTATCACTGTGCTCACCAGGA 1800
QY 1801 GCAAAAGGAGGGTTCTGCCAGAGACGGGCCCAAGAGACTTAAAGCTTCACTATAATGTGA 1860
Db 1801 GCAAAAGGAGGGTTCTGCCAGAGACGGGCCCAAGAGACTTAAAGCTTCACTATAATGTGA 1860
QY 1861 CTACAACTAGATTAGTGAATCCAGATCAACTGTGTAATGAATAATGTCTTATCTTCCAA 1920
Db 1861 CTACAACTAGATTAGTGAATCCAGATCAACTGTGTAATGAATAATGTCTTATCTTCCAA 1920
QY 1921 AGAAGCATGTGTGCTTTGTAACAAAAGGGTTATCACTGAAGTGTCAAAACACAGTCAGATT 1980
Db 1921 AGAAGCATGTGTGCTTTGTAACAAAAGGGTTATCACTGAAGTGTCAAAACACAGTCAGATT 1980
QY 1981 TTGGGAAAGTGACATGCAATTTGAATTTAGAAAGTGTGCCAGCTTCAAAAACCCGATGTGG 2040
Db 1981 TTGGGAAAGTGACATGCAATTTGAATTTAGAAAGTGTGCCAGCTTCAAAAACCCGATGTGG 2040
QY 2041 TGGGTATCAGGAGCGAGCGCTTAAAGCGGATGCTGCTGGGTTTACAAAAGATTAGTGGAG 2100
Db 2041 TGGGTATCAGGAGCGAGCGCTTAAAGCGGATGCTGCTGGGTTTACAAAAGATTAGTGGAG 2100
QY 2101 ACATCCTATCTAGCTCAAGGTATAATTTGATGGATTCTTCCATCTGCGCGGATGAGTGTG 2160
Db 2101 ACATCCTATCTAGCTCAAGGTATAATTTGATGGATTCTTCCATCTGCGCGGATGAGTGTG 2160
QY 2161 GGTGTGATACAGGCTACATAAAGACTGTTATGATGCTGTTTAAAGTTCAATTGGAA 2220
Db 2161 GGTGTGATACAGGCTACATAAAGACTGTTATGATGCTGTTTAAAGTTCAATTGGAA 2220
QY 2221 CTACCAACTTGTCTTAAAGAGCTATCTTAAGACCAATATCTCTTGTGTTTAAACAAA 2280
Db 2221 CTACCAACTTGTCTTAAAGAGCTATCTTAAGACCAATATCTCTTGTGTTTAAACAAA 2280
QY 2281 GATATTATTTGTGTATGAATCTTAATCAAGCCCATCTGTCAATATGTTACTGTCTTTT 2340
Db 2281 GATATTATTTGTGTATGAATCTTAATCAAGCCCATCTGTCAATATGTTACTGTCTTTT 2340
QY 2341 TAATCATGTGTTTGTATATTAATTAATTTGTTGACTTTCTTAGATTCATTCATATGTG 2400
Db 2341 TAATCATGTGTTTGTATATTAATTAATTTGTTGACTTTCTTAGATTCATTCATATGTG 2400
QY 2401 AATGTAAGCTCTTAATCATGTCTTTGTAATGTGTAATTTCTTCTGTAATAAACCAT 2460
Db 2401 AATGTAAGCTCTTAATCATGTCTTTGTAATGTGTAATTTCTTCTGTAATAAACCAT 2460
QY 2461 TTGTGAATAT 2470
Db 2461 TTGTGAATAT 2470
```

RESULT 8  
ADM72650

ID ADM72650 standard; cDNA; 2470 BP.

XX AC ADM72650;

XX DT 17-JUN-2004 (first entry)

XX DE Human TASK110 polypeptide encoding cDNA (clone DNA255289).

XX KW TASK; tumour-associated kinase; cytosolic; tumour;

KW cell proliferative disorder; cancer; transgenic; chromosome identification; tissue typing; human; TASK110; gene; ss.  
XX Homo sapiens.

XX Key Location/Qualifiers  
FT CDS 171..2126  
FT /\*tag= a  
FT /product= "TASK110"

WO2004024063-A2.

PN 25-MAR-2004.

XX 05-SEP-2003; 2003WO-US027886.

XX 11-SEP-2002; 2002US-0410166P.

XX (GETH ) GENENTECH INC.

XX Davis DP, Desauvage FJ, Wood WI, Zhang Z;

PI WPI; 2004-282984/26.

DR P-PSDB; ADM72651.

XX New tumor-associated kinase nucleic acids and polypeptides, useful as hybridization probes for isolating full length TASK DNA, for generating transgenic animals, in chromosome identification, or for tissue typing.

XX Claim 1; Fig 1; 140pp; English.

XX The invention relates to tumour-associated kinase (TASK) polypeptides (II) and encoding polynucleotides. An antibody, oligopeptide (siRNA) or organic molecule that binds to (II) is useful for treating a mammal having a tumour comprising cells expressing (II). Antagonists of TASK are useful for treating or preventing a cell proliferative disorder (e.g. cancer) associated with increased expression or activity of (II). The TASK polynucleotides and polypeptides may be used as hybridization probes for isolating full length TASK DNA, for generating transgenic animals, in chromosome identification, or for tissue typing. The present sequence represents a cDNA encoding a human TASK110 polypeptide.

XX Sequence 2470 BP; 767 A; 488 C; 536 G; 679 T; 0 U; 0 Other;

Query Match 100.0%; Score 2470; DB 12; Length 2470;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 2470; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTGGCGGCGGAGCGGCCCAACCCGCGATCGAAAAGATTCTTAGGAACCGCGTACCA 60

Db 1 TTGGCGGCGGAGCGGCCCAACCCGCGATCGAAAAGATTCTTAGGAACCGCGTACCA 60

QY 61 GCCGCTCTCAGACAGCAGCGCCCTGCTCTCTGCGGCGCGCTCAGCCGTGCC 120

Db 61 GCCGCTCTCAGACAGCAGCGCCCTGCTCTCTGCGGCGCGCTCAGCCGTGCC 120

QY 121 TCCGCCCTCAGGTTCTTTTCTTAATCCAAATAAATTGCAAGAGGACTATGAAGATT 180

Db 121 TCCGCCCTCAGGTTCTTTTCTTAATCCAAATAAATTGCAAGAGGACTATGAAGATT 180

QY 181 ATGATGAATCTCTCAATATTAATGAATTAATGAATTAATGGGACAGGTGCTTTGCAA 240

Db 181 ATGATGAATCTCTCAATATTAATGAATTAATGAATTAATGGGACAGGTGCTTTGCAA 240

QY 241 AGGTCAAACTTGCCTGCCATATCTTACTGAGAGATGTTAGTATATAAATCATGGATA 300

Db 241 AGGTCAAACTTGCCTGCCATATCTTACTGAGAGATGTTAGTATATAAATCATGGATA 300

QY 301 AAAACACACTAGGAGTGAATTTGCCCGGATCAAAACCGAGATTGAGGCTTTGAAGAACC 360

XX 301 AAAACACACTAGGAGTGAATTTGCCCGGATCAAAACCGAGATTGAGGCTTTGAAGAACC 360

XX 361 TGAGACATCAGCATATATGTCAACTCTACCTAGTGTCTAGACAGCAACAAATATTCA 420

[illegible]

ID ADW72216 standard; cDNA; 2470 BP.  
 AC ADW72216;  
 DT 17-JUN-2004 (first entry)  
 DE Human TASK110 polypeptide encoding cDNA (clone id DNA255289).  
 XX  
 KW TASK; tumour-associated kinase; cytostatic; tumour antigen;  
 KW cell proliferative disorder; cancer; transgenic; human; gene; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 XX Key Location/Qualifiers  
 FH 171..2126  
 FT CDS /tag= a  
 FT /product= "TASK110"  
 XX  
 PN WO2004024064-A2.  
 XX  
 PD 25-MAR-2004.  
 XX  
 PF 05-SEP-2003; 2003WO-US027894.  
 XX  
 PR 11-SEP-2002; 2002US-0410166P.  
 XX  
 PA (GETH ) GENENTECH INC.  
 XX  
 PI Desauvage FJ, Wood WI, Zhang Z;  
 XX  
 DR WPI; 2004-282985/26.  
 DR P-PSDB; ADM72217.  
 XX  
 XX New tumor-associated kinase nucleic acids and polypeptides, useful as  
 PT hybridization probes for isolating full length TASK DNA, for generating  
 PT transgenic animals, in chromosome identification, or for tissue typing.  
 XX  
 PS Claim 2; SEQ ID NO 21; 163pp; English.  
 XX  
 CC The invention relates to new isolated tumour-associated kinase (TASK)  
 CC nucleic acid molecules and encoded polypeptides. Cytostatic. The  
 CC antibody, oligopeptide or organic molecule that binds to the TASK  
 CC polypeptide are useful for treating a mammal having a tumour comprising  
 CC cells expressing the polypeptide. Antagonists of TASK are useful for  
 CC treating or preventing a cell proliferative disorder (e.g. cancer)  
 CC associated with increased expression or activity of TASK polypeptide. The  
 CC TASK polynucleotides and polypeptides may be used as hybridization probes  
 CC for isolating full length TASK DNA, for generating transgenic animals, in  
 CC chromosome identification, or for tissue typing. The present sequence  
 CC represents a cDNA encoding a human TASK polypeptide.  
 XX  
 SQ Sequence 2470 BP; 767 A; 488 C; 536 G; 679 T; 0 U; 0 Other;  
 Query Match 100.0%; Score 2470; DB 12; Length 2470;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 2470; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 TTGGGGGGGGGAGCGGCCACAAACCGGCGATCGAAAGATTCTTAGGAACCGCGTACCA 60  
 DB 1 TTGGGGGGGGGAGCGGCCACAAACCGGCGATCGAAAGATTCTTAGGAACCGCGTACCA 60  
 QY 61 GCCCGCTCTCTCAGGACAGCAGCGCCCTGTCTCTGTGCGGCGCGCTCAGCCGCTGCC 120  
 DB 61 GCCCGCTCTCTCAGGACAGCAGCGCCCTGTCTCTGTGCGGCGCGCTCAGCCGCTGCC 120  
 QY 121 TCGGCGCTCTCTCAGGACAGCAGCGCCCTGTCTCTGTGCGGCGCGCTCAGCGGATTT 180  
 DB 121 TCGGCGCTCTCTCAGGACAGCAGCGCCCTGTCTCTGTGCGGCGCGCTCAGCGGATTT 180  
 QY 181 ATGATGAACCTCTCAAAATTAATGAATACATGAACATTAATGCGGACAGTGGCTTTGCAA 240  
 DB 181 ATGATGAACCTCTCAAAATTAATGAATACATGAACATTAATGCGGACAGTGGCTTTGCAA 240

QY 241 AGGTCAAACTTGCCTGCCCATATCTCTTACTGGAGAGATGGTAGCTATAAAATCATGGATA 300  
 DB |||||  
 QY 241 AGGTCAAACTTGCCTGCCCATATCTCTTACTGGAGAGATGGTAGCTATAAAATCATGGATA 300  
 DB |||||  
 QY 301 AAAACACACTAGGGAGTGAATTTGCCCGGATCAAAACCGAGATTGAGGCTTTGAGAACC 360  
 DB |||||  
 QY 301 AAAACACACTAGGGAGTGAATTTGCCCGGATCAAAACCGAGATTGAGGCTTTGAGAACC 360  
 DB |||||  
 QY 361 TGAGACATCAGCATATATGTCAACTCTACCATGTGCTAGAGACAGCAAAATATTCA 420  
 DB |||||  
 QY 361 TGAGACATCAGCATATATGTCAACTCTACCATGTGCTAGAGACAGCAAAATATTCA 420  
 DB |||||  
 QY 421 TGGTTCTTGAGTACTGCCCTGGAGAGAGCTGTTTGAATATAATTTCCAGGATCGCC 480  
 DB |||||  
 QY 421 TGGTTCTTGAGTACTGCCCTGGAGAGAGCTGTTTGAATATAATTTCCAGGATCGCC 480  
 DB |||||  
 QY 481 TGTCAAGAGAGAGACCCGGGTTGCTTCCGTCAGATAGTATCTGCTGTTGCTTATGTC 540  
 DB |||||  
 QY 481 TGTCAAGAGAGAGACCCGGGTTGCTTCCGTCAGATAGTATCTGCTGTTGCTTATGTC 540  
 DB |||||  
 QY 541 ACAGCCAGGGCTATGCTCACAGGACCTCAAGCCAGAAAATTTGCTGTTGATGAATATC 600  
 DB |||||  
 QY 541 ACAGCCAGGGCTATGCTCACAGGACCTCAAGCCAGAAAATTTGCTGTTGATGAATATC 600  
 DB |||||  
 QY 601 ATAAATTAAGCTGATGACTTTTGGTCTCTGTGCAAAACCCCAAGGTAACAGGATTAAC 660  
 DB |||||  
 QY 601 ATAAATTAAGCTGATGACTTTTGGTCTCTGTGCAAAACCCCAAGGTAACAGGATTAAC 660  
 DB |||||  
 QY 661 ATCTACAGACATGCTGTGGAGTCTGGCTTATGCAAGACCTGAGTTAATACAGGCAAT 720  
 DB |||||  
 QY 661 ATCTACAGACATGCTGTGGAGTCTGGCTTATGCAAGACCTGAGTTAATACAGGCAAT 720  
 DB |||||  
 QY 721 CATATCTTGGATCAGAGGAGATGTTTGGAGCATGGGCATCTGTTATATGTTCTTATGT 780  
 DB |||||  
 QY 721 CATATCTTGGATCAGAGGAGATGTTTGGAGCATGGGCATCTGTTATATGTTCTTATGT 780  
 DB |||||  
 QY 781 GTGGAATTTCTACCATTTGATGATTAATGTAATGGCTTTTATACAAAGAGATATGAGAG 840  
 DB |||||  
 QY 781 GTGGAATTTCTACCATTTGATGATTAATGTAATGGCTTTTATACAAAGAGATATGAGAG 840  
 DB |||||  
 QY 841 GAAATATGATGTTCCCAAGTGGCTCTCCAGTAGGATCTGCTTCTTCAACAAATGC 900  
 DB |||||  
 QY 841 GAAATATGATGTTCCCAAGTGGCTCTCTCCAGTAGGATCTGCTTCTTCAACAAATGC 900  
 DB |||||  
 QY 901 TGCAGGTGGAACCCAAAGAAACGGATTCTTATGAAAATCTATTGAACCATCCCTGGATCA 960  
 DB |||||  
 QY 901 TGCAGGTGGAACCCAAAGAAACGGATTCTTATGAAAATCTATTGAACCATCCCTGGATCA 960  
 DB |||||  
 QY 961 TGAAGATTACAACTATCTCTGTGAGTGCAAGCAAGATCTTTTATTCACCTCGATG 1020  
 DB |||||  
 QY 961 TGAAGATTACAACTATCTCTGTGAGTGCAAGCAAGATCTTTTATTCACCTCGATG 1020  
 DB |||||  
 QY 1021 ATGATGGGTAAACGAACCTTCTGTACATCAAGAAACCAAGCAACCAATGGAGGATT 1080  
 DB |||||  
 QY 1021 ATGATGGGTAAACGAACCTTCTGTACATCAAGAAACCAAGCAACCAATGGAGGATT 1080  
 DB |||||  
 QY 1081 TAAATTTCACTGTGGCAGTATGATCACTCAGCGCTACCTATCTTCTGTCTTAGCAAGA 1140  
 DB |||||  
 QY 1081 TAAATTTCACTGTGGCAGTATGATCACTCAGCGCTACCTATCTTCTGTCTTAGCAAGA 1140  
 DB |||||  
 QY 1141 AGGCTCGGGGAAAACCAAGTCTGTTTAAAGGCTTTCTTTCTCTGTGTGGAAGAAGCAGTG 1200  
 DB |||||  
 QY 1141 AGGCTCGGGGAAAACCAAGTCTGTTTAAAGGCTTTCTTTCTCTGTGTGGAAGAAGCAGTG 1200  
 DB |||||  
 QY 1201 CTACCCCATTCACAGACATCAAGTCAAAATTTGAGTCTGGAAGATCTGACCGCAAGTG 1260  
 DB |||||  
 QY 1201 CTACCCCATTCACAGACATCAAGTCAAAATTTGAGTCTGGAAGATCTGACCGCAAGTG 1260  
 DB |||||  
 QY 1261 ATAAATTTATGTGGCGGATTAATAGACTATGATTGGTGTGAAGATGATTATCAACAG 1320  
 DB |||||  
 QY 1261 ATAAATTTATGTGGCGGATTAATAGACTATGATTGGTGTGAAGATGATTATCAACAG 1320  
 DB |||||  
 QY 1321 GTGCTGCTACTCCCGCAACATCAAGTTTACAAAGTACTGGACAGAAATCAAAATGGGTGG 1380

1321 GTGCTGCTACTCCCGGAAACATCACAGTTTACCAAGTACTGGACAGAAATCAAAATGGGGTGG 1380  
1381 AATCTAAATCAATTAATCCAGCCTTATGAGACACCTCCAAATAATTAAGAACAAG 1440  
1381 AATCTAAATCAATTAATCCAGCCTTATGAGACACCTCCAAATAATTAAGAACAAG 1440  
1441 AAAATGTATATCTCTCTAGTCTGTGTAAGAATGAAGTACTTTATGTCTTCCTGAGC 1500  
1441 AAAATGTATATCTCTCTAGTCTGTGTAAGAATGAAGTACTTTATGTCTTCCTGAGC 1500  
1501 CAAAGACTCCAGTTAATAAGAACCCAGCATAGAGAGAAATATCTCATCGCCAAATCGTT 1560  
1501 CAAAGACTCCAGTTAATAAGAACCCAGCATAGAGAGAAATATCTCATCGCCAAATCGTT 1560  
1561 ACATACACCTCAAAAGCTAGAAACCCAGTCCCTGAAAGAACTCAATTAATAATCCAG 1620  
1561 ACATACACCTCAAAAGCTAGAAACCCAGTCCCTGAAAGAACTCAATTAATAATCCAG 1620  
1621 TAAATTTCAACAGGAACAGACAAGTTAATGACAGGTGTCTATTAGCCCTGAGAGCGGTGCC 1680  
1621 TAAATTTCAACAGGAACAGACAAGTTAATGACAGGTGTCTATTAGCCCTGAGAGCGGTGCC 1680  
1681 GCTCAGTGGAAATGGATCTCAACCAAGCACATATGAGAGGAGCTCCAAAAGAAAGGGAG 1740  
1681 GCTCAGTGGAAATGGATCTCAACCAAGCACATATGAGAGGAGCTCCAAAAGAAAGGGAG 1740  
1741 CCAAAGTGTGGAGGCTTGAAGGGGTTGATAGAGTATCACTGTGCTCACCAGGA 1800  
1741 CCAAAGTGTGGAGGCTTGAAGGGGTTGATAGAGTATCACTGTGCTCACCAGGA 1800  
1801 GCAAAAGGAGGGTCTCGCCAGAGACGGGCCAGAGACTTAAAGCTTCACTATAATGTCA 1860  
1801 GCAAAAGGAGGGTCTCGCCAGAGACGGGCCAGAGACTTAAAGCTTCACTATAATGTCA 1860  
1861 CTAACAATAGATTAGTGAATCCAGATCAACTGTGTGAATGAATAATGTCTATTTCCAA 1920  
1861 CTAACAATAGATTAGTGAATCCAGATCAACTGTGTGAATGAATAATGTCTATTTCCAA 1920  
1921 AGAAGCATGTGTGATCTTGTACAAAGGGTTATACACTGAAGTGTCAAAACACAGTCAAGT 1980  
1921 AGAAGCATGTGTGATCTTGTACAAAGGGTTATACACTGAAGTGTCAAAACACAGTCAAGT 1980  
1981 TTGGGAAAGTGACAAATGCAATTTGAATAGAGTGTCCAGCTTCAAAAACCCGATGTGG 2040  
1981 TTGGGAAAGTGACAAATGCAATTTGAATAGAGTGTCCAGCTTCAAAAACCCGATGTGG 2040  
2041 TGGGTATCAGGAGGCGCGCTTAAAGGCGATCCCTGGGTTTACAAAAGATTAGTGGAG 2100  
2041 TGGGTATCAGGAGGCGCGCTTAAAGGCGATCCCTGGGTTTACAAAAGATTAGTGGAG 2100  
2101 ACATCTATCTAGCTGCAAGGTATTAATGATGATTCCTCCATCTGCGGATGAGTGG 2160  
2101 ACATCTATCTAGCTGCAAGGTATTAATGATGATTCCTCCATCTGCGGATGAGTGG 2160  
2161 GGTGTGATACAGGCTACATAAAGACTGTGTTATGATCGCTTTGATTTTAAAGTTCATTGGAA 2220  
2161 GGTGTGATACAGGCTACATAAAGACTGTGTTATGATCGCTTTGATTTTAAAGTTCATTGGAA 2220  
2221 CTAACCAACTTGTCTTAAAGAGCTATCTTAAGACCAATATCTCTTTGTTTTTAAACAAA 2280  
2221 CTAACCAACTTGTCTTAAAGAGCTATCTTAAGACCAATATCTCTTTGTTTTTAAACAAA 2280  
2281 GATATTTTGTGTATGATTAATCAATCAAGCCCATCTGTCATATGTTACTGTCTTTT 2340  
2281 GATATTTTGTGTATGATTAATCAATCAAGCCCATCTGTCATATGTTACTGTCTTTT 2340  
2341 TAATCATGTGTTTGTATATTAATATTTGACCTTCTTAGATTCACCTTCCATATGTG 2400  
2341 TAATCATGTGTTTGTATATTAATATTTGACCTTCTTAGATTCACCTTCCATATGTG 2400  
2401 AATGTAAGCTTTAACTATGCTCTTTTGTAAATGTGTAATTTCTTCTGAAATAAAACCAT 2460

2401 AATGTAAGCTTTAACTATGCTCTTTTGTAAATGTGTAATTTCTTCTGAAATAAAACCAT 2460  
2461 TTGTGAATAT 2470  
2461 TTGTGAATAT 2470  
2461 TTGTGAATAT 2470  
RESULT 10  
ADN06035  
ID ADN06035 standard; cdNA; 2470 BP.  
XX  
AC ADN06035;  
XX  
XX  
DT 01-JUL-2004 (first entry)  
XX  
DE Antipsoriatic cdNA sequence #1256.  
XX  
XX ds; gene; antipsoriatic; gene therapy; psoriasis; diagnosis.  
KW  
XX Homo sapiens.  
OS  
XX WO2004028479-A2.  
PN  
XX  
XX 08-APR-2004.  
PD  
XX  
XX 25-SEP-2003; 2003WO-US030907.  
PF  
XX  
PR 25-SEP-2002; 2002US-0414006P.  
XX  
XX (GETH ) GENENTECH INC.  
PA  
XX  
PI Bodary S, Clark H, Jackman J, Schoenfeld J, Williams PM, Wood WI;  
PI Wu TD;  
XX  
XX WPI; 2004-305105/28.  
DR P-PSDB; ADN06036.  
XX  
XX New PRO nucleic acid or polypeptide, useful for preparing a  
PT pharmaceutical composition for diagnosing or treating psoriasis in a  
PT mammal.  
XX  
XX Claim 1; SEQ ID NO 2430; 3069pp; English.  
XX  
XX The invention relates to novel polynucleotide and polypeptides for  
CC treating psoriasis or a sequence having at least 80% identity to the  
CC above sequences. The nucleic acid is useful for preparing a composition  
CC for diagnosing or treating psoriasis in a mammal. This sequence  
CC corresponds to one of the polynucleotides of the invention.  
XX  
XX Sequence 2470 BP; 767 A; 488 C; 536 G; 679 T; 0 U; 0 Other;  
Query Match 100.0%; Score 2470; DB 12; Length 2470;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2470; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 TTGGGGGGCGGAGCGGCCACCAACCGCGGATCGAAAGATTCTTAGGAACCGCGTACCA 60  
DB 1 TTGGGGGGCGGAGCGGCCACCAACCGCGGATCGAAAGATTCTTAGGAACCGCGTACCA 60  
QY 61 GCGGGGTCTCTCAGACACAGCGCCCTGTCTTCTGTGGGGCGCGCTCAGCCGTGCC 120  
DB 61 GCGGGGTCTCTCAGACACAGCGCCCTGTCTTCTGTGGGGCGCGCTCAGCCGTGCC 120  
QY 121 TCCGCCCTCAGTCTCTTTTCTTAATTCCAAATAAATTCGAAGAGACTATGAAGAATT 180  
DB 121 TCCGCCCTCAGTCTCTTTTCTTAATTCCAAATAAATTCGAAGAGACTATGAAGAATT 180  
QY 181 ATGATGAACCTCTCAAATAATTAATGAAATACATGAAACTATGGAACAGGTGCGTTCGAA 240  
DB 181 ATGATGAACCTCTCAAATAATTAATGAAATACATGAAACTATGGAACAGGTGCGTTCGAA 240  
QY 241 AGGTCAAACCTTGGCTGCCATATCTTACTGGAGAGATGTTAGCTATAAATAATCATCGATA 300

Db 241 AGGTCAAACTTGCTGCCATATCTTACTGGAGAGATGGTAGCTATAAAAAATCATGGATA 300  
Qy 301 AAAACACACTAGGAGGTGATTTGCCCGCGATCAAAAACGGAGATGAGGCGCTTGAAGAAC 360  
Db 301 AAAACACACTAGGAGGTGATTTGCCCGCGATCAAAAACGGAGATGAGGCGCTTGAAGAAC 360  
Qy 361 TGAGACATCAGCATATATGTCAACTCTACCATGTGCTAGAGACAGCCAAACAAATATTC 420  
Db 361 TGAGACATCAGCATATATGTCAACTCTACCATGTGCTAGAGACAGCCAAACAAATATTC 420  
Qy 421 TGGTTCTTGAGTACTGCCCTGGAGGAGAGCTGTTTGACTATATAAATTTCCAGGATCGCC 480  
Db 421 TGGTTCTTGAGTACTGCCCTGGAGGAGAGCTGTTTGACTATATAAATTTCCAGGATCGCC 480  
Qy 481 TGTCAAGAGGAGAGACCGGGTTGCTTCGGTCAGATAGTATCTGCTGTGCTTATGTGC 540  
Db 481 TGTCAAGAGGAGAGACCGGGTTGCTTCGGTCAGATAGTATCTGCTGTGCTTATGTGC 540  
Qy 541 ACAGCCAGGGCTATGCTCAAGGGACCTCAAGCCAGAAAATTTGCTGTTTGATGAATATC 600  
Db 541 ACAGCCAGGGCTATGCTCAAGGGACCTCAAGCCAGAAAATTTGCTGTTTGATGAATATC 600  
Qy 601 ATAAATTAAGCTGATGACTTTGGTCTCTGTGCAAAACCCAGGGGTAAACAAGGATTACC 660  
Db 601 ATAAATTAAGCTGATGACTTTGGTCTCTGTGCAAAACCCAGGGGTAAACAAGGATTACC 660  
Qy 661 ATCTACAGACATGCTGTGGAGGCTCGGCTTATGACGACCTGAGTAAATACAAAGGCAAT 720  
Db 661 ATCTACAGACATGCTGTGGAGGCTCGGCTTATGACGACCTGAGTAAATACAAAGGCAAT 720  
Qy 721 CATATCTTGATCAGAGGAGATGTTTGAGCATGCGGCTACTGTTTATATGTTCTTATGT 780  
Db 721 CATATCTTGATCAGAGGAGATGTTTGAGCATGCGGCTACTGTTTATATGTTCTTATGT 780  
Qy 781 GTGGATTTCTACCATTTGATGATGATAATGTAATGGCTTTATACAAAGAGATTTATGAGAG 840  
Db 781 GTGGATTTCTACCATTTGATGATGATAATGTAATGGCTTTATACAAAGAGATTTATGAGAG 840  
Qy 841 GAAAATATGATGTTCCAAAGTGCTCTCCAGTAGCATCTGCTTCTTCAACAAATGC 900  
Db 841 GAAAATATGATGTTCCAAAGTGCTCTCCAGTAGCATCTGCTTCTTCAACAAATGC 900  
Qy 901 TGCAGGTGACCCAAAGAACGGATTTCTATGAAAATCTATGAAAATCTATTTGAACCATCCCTGGATCA 960  
Db 901 TGCAGGTGACCCAAAGAACGGATTTCTATGAAAATCTATGAAAATCTATTTGAACCATCCCTGGATCA 960  
Qy 961 TGCAGGATTAACAATCTATCTGTTGAGTGGCAAGCAAGAAATCCTTTTATTCACCTCGATG 1020  
Db 961 TGCAGGATTAACAATCTATCTGTTGAGTGGCAAGCAAGAAATCCTTTTATTCACCTCGATG 1020  
Qy 1021 ATGATTTGCGTAAACAGAACTTTCTGTATCATCAGAAAACAAAGGCAAAACAATGGAGGATT 1080  
Db 1021 ATGATTTGCGTAAACAGAACTTTCTGTATCATCAGAAAACAAAGGCAAAACAATGGAGGATT 1080  
Qy 1081 TAAATTTCACTGTCGAGTATGATACCTCAGCGCTACCTATCTTCTGCTTCTAGCCAAAGA 1140  
Db 1081 TAAATTTCACTGTCGAGTATGATACCTCAGCGCTACCTATCTTCTGCTTCTAGCCAAAGA 1140  
Qy 1141 AGGCTCGGGGAAAACAGTTCGTTTAAAGGCTTTCTTTCTTCTGTCGACCAAGCCAGTG 1200  
Db 1141 AGGCTCGGGGAAAACAGTTCGTTTAAAGGCTTTCTTTCTTCTGTCGACCAAGCCAGTG 1200  
Qy 1201 CTACCCCAATTCACAGACATCAAGTCAAAATTTGGAGTCTGGAAGATGTGACCGCAAGTG 1260  
Db 1201 CTACCCCAATTCACAGACATCAAGTCAAAATTTGGAGTCTGGAAGATGTGACCGCAAGTG 1260  
Qy 1261 ATAAAAATTTATGTGGGGGATTAATAGACTATGATTTGGTGAAGATGATTTATCAACAG 1320  
Db 1261 ATAAAAATTTATGTGGGGGATTAATAGACTATGATTTGGTGAAGATGATTTATCAACAG 1320  
Qy 1321 GTGCTGTACTCCCGAAACATCACAGTTTATACCAAGTCTGGACAGAAATCAAAATGGGTGG 1380  
Db 1321 GTGCTGTACTCCCGAAACATCACAGTTTATACCAAGTCTGGACAGAAATCAAAATGGGTGG 1380

Qy 1381 AATCTAAATCATTAATCTCAGGCTTATGAGAACCACTGCAAAATAAATTAAGAAACAAAG 1440  
Db 1381 AATCTAAATCATTAATCTCAGGCTTATGAGAACCACTGCAAAATAAATTAAGAAACAAAG 1440  
Qy 1441 AAAATGTATATACTCCTTAAGTCTGCTGTAAGAAAGTAAGAGTACTTTATGTTTCTGAGC 1500  
Db 1441 AAAATGTATATACTCCTTAAGTCTGCTGTAAGAAAGTAAGAGTACTTTATGTTTCTGAGC 1500  
Qy 1501 CAAAGACTCCAGTTAATTAAGAACCAAGCAGATAAGAGAGAAATATCTACTACGCCAAATCGTT 1560  
Db 1501 CAAAGACTCCAGTTAATTAAGAACCAAGCAGATAAGAGAGAAATATCTACTACGCCAAATCGTT 1560  
Qy 1561 ACATACACCTTCAAAGCTAGAAACCAAGTCTGAAAGAAACTCAATTAATAATACAG 1620  
Db 1561 ACATACACCTTCAAAGCTAGAAACCAAGTCTGAAAGAAACTCAATTAATAATACAG 1620  
Qy 1621 TAAATTTCAACAGGACAGACAGCAAGTTAATGACAGGTGTCAATTAGCCCTGAGAGCGGTGCC 1680  
Db 1621 TAAATTTCAACAGGACAGACAGCAAGTTAATGACAGGTGTCAATTAGCCCTGAGAGCGGTGCC 1680  
Qy 1681 GCTCAGTGGAAATTGGATCTCAACCAAGCACAATATGAGAGAGACTCCAAAAGAAAGGGAG 1740  
Db 1681 GCTCAGTGGAAATTGGATCTCAACCAAGCACAATATGAGAGAGACTCCAAAAGAAAGGGAG 1740  
Qy 1741 CAAAAGTGTGGAGCCCTTGAAAGGGGTGGATAAGGTTATCACTGTGCTCACCAGGA 1800  
Db 1741 CAAAAGTGTGGAGCCCTTGAAAGGGGTGGATAAGGTTATCACTGTGCTCACCAGGA 1800  
Qy 1801 GCAAAAGGAGGGTTCGCCAGAGAGCGGCCAGAGACTAAAGCTTCACTATATATGTA 1860  
Db 1801 GCAAAAGGAGGGTTCGCCAGAGAGCGGCCAGAGACTAAAGCTTCACTATATATGTA 1860  
Qy 1861 CTAACTAGATTAAGTGAATCCAGATCAACTGTTGAATGAATAATGTCTATCTTCCAA 1920  
Db 1861 CTAACTAGATTAAGTGAATCCAGATCAACTGTTGAATGAATAATGTCTATCTTCCAA 1920  
Qy 1921 AGAAGCATGTTGACTTTGTGTAACAAAAGGGTTATACACTGAAGTGTCAAAACACAGTCA 1980  
Db 1921 AGAAGCATGTTGACTTTGTGTAACAAAAGGGTTATACACTGAAGTGTCAAAACACAGTCA 1980  
Qy 1981 TTGGGAAAGTGCATGCAATTTGAAATTTAGAAATTTAGAAATTTAGAAATTTAGTGG 2040  
Db 1981 TTGGGAAAGTGCATGCAATTTGAAATTTAGAAATTTAGAAATTTAGAAATTTAGTGG 2040  
Qy 2041 TGGGTATCAGGAGGCGGCTTAAAGGCGATGCTGCTGTTTACAAAAGATTAAGTGGAG 2100  
Db 2041 TGGGTATCAGGAGGCGGCTTAAAGGCGATGCTGCTGTTTACAAAAGATTAAGTGGAG 2100  
Qy 2101 ACATCTATCTAGCTGCAAGGTATAATTTAGTGGATTTCTTCCATCTGCGGATGAGTGTG 2160  
Db 2101 ACATCTATCTAGCTGCAAGGTATAATTTAGTGGATTTCTTCCATCTGCGGATGAGTGTG 2160  
Qy 2161 GGTGTGATACAGCTTACATAAAGACTGTTATGATCGCTTTGATTTTAAAGTTCATTGGAA 2220  
Db 2161 GGTGTGATACAGCTTACATAAAGACTGTTATGATCGCTTTGATTTTAAAGTTCATTGGAA 2220  
Qy 2221 CTACCAACTGTTTCTAAAGAGCTATCTTAAGACCAATATCTCTTTGTTTAAACAAA 2280  
Db 2221 CTACCAACTGTTTCTAAAGAGCTATCTTAAGACCAATATCTCTTTGTTTAAACAAA 2280  
Qy 2281 GATATTTATTTGTGATGAATCTAAATCAAGCCCATCTGTCAATTAATGTTACTGTCTTTT 2340  
Db 2281 GATATTTATTTGTGATGAATCTAAATCAAGCCCATCTGTCAATTAATGTTACTGTCTTTT 2340  
Qy 2341 TAAATCATGTTTGTATATATAATTAATGTTGACTTTCTAGATTTCACTTCCATATGTG 2400  
Db 2341 TAAATCATGTTTGTATATATAATTAATGTTGACTTTCTAGATTTCACTTCCATATGTG 2400  
Qy 2401 AATGTAAGCTCTTAACATATGCTCTTTGTAATGTAATTTCTTCTGAAATFAAACCAT 2460  
Db 2401 AATGTAAGCTCTTAACATATGCTCTTTGTAATGTAATTTCTTCTGAAATFAAACCAT 2460





```
Db 1261 ATAAAAATTAATGTCGGCGGATTAATAGACTATGATTTGGTGTGAAGATGATTTATCAACAG 1320
QY GTGCTGTACTCCCGGAAATCACAGTTTACCAAGTACTGGACAGAAATCAAAATGGGGTGG 1380
Db GTGCTGTACTCCCGGAAATCACAGTTTACCAAGTACTGGACAGAAATCAAAATGGGGTGG 1380
QY AATCTAAATCAATTAATCTCCAGCCTTATGAGAAACACTGTCGAAATTAATAAGAAACAAAG 1440
Db AATCTAAATCAATTAATCTCCAGCCTTATGAGAAACACTGTCGAAATTAATAAGAAACAAAG 1440
QY AAAATGTATATACCTCTAAGTCTGCTGAAGAAATGAAGAGTACTTTATGTTTCTCGAGC 1500
Db AAAATGTATATACCTCTAAGTCTGCTGAAGAAATGAAGAGTACTTTATGTTTCTCGAGC 1500
QY CAAAGACTCCAGTTAATAAGAACCCAGCAGATAGAGAGAAATCTCACTACGCCAAATCGTT 1560
Db CAAAGACTCCAGTTAATAAGAACCCAGCAGATAGAGAGAAATCTCACTACGCCAAATCGTT 1560
QY AACTACACCTCAAAAGCTAGAAACAGTGCCTGAAAGAACTCAATTAATAATACCAG 1620
Db AACTACACCTCAAAAGCTAGAAACAGTGCCTGAAAGAACTCAATTAATAATACCAG 1620
QY TAAATTTCAACAGGAAACAGCAAGTTAATGACAGGTCTATTAGCCTGAGAGCGGTGCC 1680
Db TAAATTTCAACAGGAAACAGCAAGTTAATGACAGGTCTATTAGCCTGAGAGCGGTGCC 1680
QY GCTCAGTGAATTTGGATCTCAACCAAGCACATATGAGAGAGACTCCAAAAGAAAGGGAG 1740
Db GCTCAGTGAATTTGGATCTCAACCAAGCACATATGAGAGAGACTCCAAAAGAAAGGGAG 1740
QY CCAAAGTGTGTTGGAGCCTTGAAGGGGTTGGATTAAGTTTACCTGCTCACCAGGA 1800
Db CCAAAGTGTGTTGGAGCCTTGAAGGGGTTGGATTAAGTTTACCTGCTCACCAGGA 1800
QY GCAAAAGGAGGGTTCTGCGAGAGAGCGGCCGAGAGACTTAAAGCTTCACTATATGTGA 1860
Db GCAAAAGGAGGGTTCTGCGAGAGAGCGGCCGAGAGACTTAAAGCTTCACTATATGTGA 1860
QY CTACAACTAGATTAGTGAATCCAGATCAACTGTGTAATGAATAATGTTCTATCTTCCAA 1920
Db CTACAACTAGATTAGTGAATCCAGATCAACTGTGTAATGAATAATGTTCTATCTTCCAA 1920
QY AGAAGCATGTTGACTTTGTACAAAAGGGTTATACACTGAAGTGTCAACACAGTCAGATT 1980
Db AGAAGCATGTTGACTTTGTACAAAAGGGTTATACACTGAAGTGTCAACACAGTCAGATT 1980
QY TTGGGAACTGACATGCAATTTGANTTAGAAGTGTGCCAGCTTCAAAACCGGATGG 2040
Db TTGGGAACTGACATGCAATTTGANTTAGAAGTGTGCCAGCTTCAAAACCGGATGG 2040
QY TTGGGAAAGTGACATGCAATTTGANTTAGAAGTGTGCCAGCTTCAAAACCGGATGG 2040
Db TTGGGAAAGTGACATGCAATTTGANTTAGAAGTGTGCCAGCTTCAAAACCGGATGG 2040
QY TGGGTATCAGGAGGCGGCTTAAGGGCGATGCCCTGGGTTTACAAAAGATTAGTGGAG 2100
Db TGGGTATCAGGAGGCGGCTTAAGGGCGATGCCCTGGGTTTACAAAAGATTAGTGGAG 2100
QY ACATCCTATCTAGCTGCAAGGTATATATCATGATTTCTTCCATCTCGCGGATGATGTG 2160
Db ACATCCTATCTAGCTGCAAGGTATATATCATGATTTCTTCCATCTCGCGGATGATGTG 2160
QY GTGTGTATACAGCTACATTAAGACTGTTATGATCGCTTTGATTTTAAAGTTCAATGGAA 2220
Db GTGTGTATACAGCTACATTAAGACTGTTATGATCGCTTTGATTTTAAAGTTCAATGGAA 2220
QY CTACCAACTTGTGTTCTTAAAGAGCTATCTTTAAGACCAATATCTCTTTGTTTTTAAACAAA 2280
Db CTACCAACTTGTGTTCTTAAAGAGCTATCTTTAAGACCAATATCTCTTTGTTTTTAAACAAA 2280
QY GATATTTATTTGTGATGATTAATCAAGGCCATCTGTCATATGTTACTGTCTTTTT 2340
Db GATATTTATTTGTGATGATTAATCAAGGCCATCTGTCATATGTTACTGTCTTTTT 2340
QY TAAATCATGTGGTTTGTATATTAATTAATGTTGACCTTCTTGAATTTCACTTCCATATGTG 2400
Db TAAATCATGTGGTTTGTATATTAATTAATGTTGACCTTCTTGAATTTCACTTCCATATGTG
```

```
Db 2341 TAAATCATGTGGTTTGTATATTAATTAATTAATTAATTAATTAATTAATTAATTAATG 2400
QY 2401 AATGTAAGCTCTTAATCATGTCTCTTTGTAATGTGTAATTTCTTTCTGAAATAAAACCAT 2460
Db 2401 AATGTAAGCTCTTAATCATGTCTCTTTGTAATGTGTAATTTCTTTCTGAAATAAAACCAT 2460
QY 2461 TTGTGAATAT 2470
Db 2461 TTGTGAATAT 2470

RESULT 12
ADR25674
ID ADR25674 standard; DNA; 2470 BP.
XX
AC ADR25674;
DT 21-OCT-2004 (first entry)
XX
DE Breast cancer prognosis marker #1535.
XX
KW ds; breast cancer; prognosis; gene expression; diagnosis.
XX
OS Homo sapiens.
XX
PN WO2004065545-A2.
XX
PD 05-AUG-2004.
XX
PF 15-JAN-2004; 2004WO-US001100.
XX
PR 15-JAN-2003; 2003US-00342887.
XX
PA (ROSE-) ROSETTA INPHARMATICS LLC.
XX
PA (NECA-) NETHERLANDS CANCER INST.
XX
PI Van't Veer LJ, He Y;
XX
WPI; 2004-593473/57.

Classifying a breast cancer patient according to prognosis comprises
determining the similarity between the level of expression of each of
five genes in a cell sample taken from patient, to control levels.

Disclosure; SEQ ID NO 1535; 226pp; English.

The invention relates to a method of classifying a breast cancer patient
according to prognosis by determining the similarity between the level of
expression of each of five genes for which markers are listed in the
specification, in a cell sample taken from the breast cancer patient, to
control levels of expression for each respective five genes to obtain a
patient similarity value. The methods are useful for classifying a breast
cancer patient according to prognosis, kits and computer program products
are useful for data analysis using the diagnostic, prognostic and
statistical methods of the invention. This sequence corresponds to a
marker used in the method of the invention.

Sequence 2470 BP; 767 A; 488 C; 536 G; 679 T; 0 U; 0 Other;

Query Match 100.0%; Score 2470; DB 13; Length 2470;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2470; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTGGCGGCGGAAGCGGCCACAAACCGCGGATCGAAAGATTTCTTAGGAACGCCGTACCA 60
Db 1 TTGGCGGCGGAAGCGGCCACAAACCGCGGATCGAAAGATTTCTTAGGAACGCCGTACCA 60
QY 61 GCCCGGTCTCTCAGGACAGCAGGCCCTGTCTCTTCTGTGCGGCGCGCTCAGCGCGTCCC 120
Db 61 GCCCGGTCTCTCAGGACAGCAGGCCCTGTCTCTTCTGTGCGGCGCGCTCAGCGCGTCCC 120
QY 121 TCCGCCCTCTCAGGTTCTTTTCTTAATTCAAATAAATTTGCAAGGAGCTATGAAGATT 180
Db 121 TCCGCCCTCTCAGGTTCTTTTCTTAATTCAAATAAATTTGCAAGGAGCTATGAAGATT 180
```

Db 121 TCCGCCCTCAGGTTCTTTTCTAATTTCCAAATAAACTTTCGAAGAGCACTATGAAGAATT 180  
Qy 181 ATGATGAACCTTCTCAAAATATATGAATATACATGAATACTATGGGACAGGTGGCTTGGCAA 240  
Db 181 ATGATGACCTTCTCAAAATATATGAATATACATGAATACTATGGGACAGGTGGCTTGGCAA 240  
Qy 241 AGGTCAAACTTGGCTGCCATATCCTTTACTGAGAGATGGTAGCTATAAAAAATCATGGATA 300  
Db 241 AGGTCAAACTTGGCTGCCATATCCTTTACTGAGAGATGGTAGCTATAAAAAATCATGGATA 300  
Qy 301 AAAAAACACTAGGAGTATTTGCCCGGATCAAAAACGAGATTTGAGGCTTGAAGAACC 360  
Db 301 AAAAAACACTAGGAGTATTTGCCCGGATCAAAAACGAGATTTGAGGCTTGAAGAACC 360  
Qy 361 TGAGACATCAGCATATATGTGCAACTCTACCATGTGCTAGAGACAGCCAAACAAATATTCA 420  
Db 361 TGAGACATCAGCATATATGTGCAACTCTACCATGTGCTAGAGACAGCCAAACAAATATTCA 420  
Qy 421 TGGTTCTTGAGTACTGCCCTGGAGGAGAGCTGTTTGACTATATAAATTTCCAGGATCGCC 480  
Db 421 TGGTTCTTGAGTACTGCCCTGGAGGAGAGCTGTTTGACTATATAAATTTCCAGGATCGCC 480  
Qy 481 TGTGAGAAGAGAGACCCGGGTTGTCTTCGGTCAGATAGTATCTGCTGTTGCTTATGTGC 540  
Db 481 TGTGAGAAGAGAGACCCGGGTTGTCTTCGGTCAGATAGTATCTGCTGTTGCTTATGTGC 540  
Qy 541 ACAGCCAGGCTATGCTCAAGGAGCTCAAGCCAGAAATTTGCTGTTGATGAATATC 600  
Db 541 ACAGCCAGGCTATGCTCAAGGAGCTCAAGCCAGAAATTTGCTGTTGATGAATATC 600  
Qy 601 ATAAATTAAGCTGATTTGCTCTGTGTCAGAAACCCAAAGGCTAAACAGGATTAAC 660  
Db 601 ATAAATTAAGCTGATTTGCTCTGTGTCAGAAACCCAAAGGCTAAACAGGATTAAC 660  
Qy 661 ATCTACAGACATGCTGGGAGTCTGGCTTATGACAGCACTGATGTAATACAAAGCAAT 720  
Db 661 ATCTACAGACATGCTGGGAGTCTGGCTTATGACAGCACTGATGTAATACAAAGCAAT 720  
Qy 721 CATATCTTGATCAGAGGAGATGTTTGGAGATGGCATCTGTTATATGTTCTTTATGT 780  
Db 721 CATATCTTGATCAGAGGAGATGTTTGGAGATGGCATCTGTTATATGTTCTTTATGT 780  
Qy 781 GTGGATTTCTACCATTTGATGATAATGTAATGCTTTTATACAAAGAGATPATGAGAG 840  
Db 781 GTGGATTTCTACCATTTGATGATAATGTAATGCTTTTATACAAAGAGATPATGAGAG 840  
Qy 841 GAAAAATATGATGTTCCCAAGTGGCTCTCCCAAGTAGCATCTGCTTCTTCAACAAATGC 900  
Db 841 GAAAAATATGATGTTCCCAAGTGGCTCTCTCCCAAGTAGCATCTGCTTCTTCAACAAATGC 900  
Qy 901 TGCAGGTGGACCCAAAGAACGGATTTCTATGAAAAATCTATTGAACCATCCCTGGATCA 960  
Db 901 TGCAGGTGGACCCAAAGAACGGATTTCTATGAAAAATCTATTGAACCATCCCTGGATCA 960  
Qy 961 TGCAAGATTACAATCTCTGTTGAGTGGCAAAAGCAAGATCCTTTATTCACCTCGATG 1020  
Db 961 TGCAAGATTACAATCTCTGTTGAGTGGCAAAAGCAAGATCCTTTATTCACCTCGATG 1020  
Qy 1021 ATGATTTGCGTAAACAGAACTTCTGTACATCAAGAAACCAACAGGCAAAACAAATGGAGATT 1080  
Db 1021 ATGATTTGCGTAAACAGAACTTCTGTACATCAAGAAACCAACAGGCAAAACAAATGGAGATT 1080  
Qy 1081 TAAATTTCACTGGCAGTAGATCACCCTCAGGCTACCTATCTTCTGCTTCTAGCCAGA 1140  
Db 1081 TAAATTTCACTGGCAGTAGATCACCCTCAGGCTACCTATCTTCTGCTTCTAGCCAGA 1140  
Qy 1141 AGGCTCGGGAAAAACAGTTTCGTTTAAAGGCTTTCTTCTTCTGTTGGAACCCAGTG 1200  
Db 1141 AGGCTCGGGAAAAACAGTTTCGTTTAAAGGCTTTCTTCTTCTTCTGTTGGAACCCAGTG 1200  
Qy 1201 CTACCCCATTCACAGACATCAAGTCAAAATTTGGAGTCTGGAAGATGTGACCCGAAGTG 1260  
Db 1201 CTACCCCATTCACAGACATCAAGTCAAAATTTGGAGTCTGGAAGATGTGACCCGAAGTG 1260

Qy 1261 ATAAAAATTTATGTGCGGGATTAATAGACTATGATTTGGTGTGAAGATGATTTATCAACAG 1320  
Db 1261 ATAAAAATTTATGTGCGGGATTAATAGACTATGATTTGGTGTGAAGATGATTTATCAACAG 1320  
Qy 1321 GTGCTGCTACTCCCGAAACATCAGATTTACCAAGTACTGGACAGAAATCAAAATGGGTGG 1380  
Db 1321 GTGCTGCTACTCCCGAAACATCAGATTTACCAAGTACTGGACAGAAATCAAAATGGGTGG 1380  
Qy 1381 AATCTAATAATCATTAATCTCCAGCTTTATGCAGAAACCTGCAAAATAAATTTAAAGAAACAAAG 1440  
Db 1381 AATCTAATAATCATTAATCTCCAGCTTTATGCAGAAACCTGCAAAATAAATTTAAAGAAACAAAG 1440  
Qy 1441 AAAATGTATATACTCCTTAAGTCTGCTGTAAGAAATGAAGAGTACTTTATGTTTCTGTAGC 1500  
Db 1441 AAAATGTATATACTCCTTAAGTCTGCTGTAAGAAATGAAGAGTACTTTATGTTTCTGTAGC 1500  
Qy 1501 CAAAGACTCCAGTTAATAAGAACCCAGTATGAGAGAGAAATCTCACTACGCCCAATCGTT 1560  
Db 1501 CAAAGACTCCAGTTAATAAGAACCCAGTATGAGAGAGAAATCTCACTACGCCCAATCGTT 1560  
Qy 1561 ACACCTACACCTCAAAAGCTAGAAACCCAGTCTGAAAGAAACTCCAAATTAATAATACAG 1620  
Db 1561 ACACCTACACCTCAAAAGCTAGAAACCCAGTCTGAAAGAAACTCCAAATTAATAATACAG 1620  
Qy 1621 TAAATTTCAACAGGAAACAGACAAAGTTAATGACAGAGTGTCAATTAGCCCTGAGAGCGGTGCC 1680  
Db 1621 TAAATTTCAACAGGAAACAGACAAAGTTAATGACAGAGTGTCAATTAGCCCTGAGAGCGGTGCC 1680  
Qy 1681 GCTCAGTGAATTTGATCTCAACCAAGCACATATGAGAGAGACTCCAAAGAAAGAGGAG 1740  
Db 1681 GCTCAGTGAATTTGATCTCAACCAAGCACATATGAGAGAGACTCCAAAGAAAGAGGAG 1740  
Qy 1741 CCAAGTGTTCGGAGCCCTTGAAGGGGTTGTAAGGTTTATCACTGTGCTCACCAGGA 1800  
Db 1741 CCAAGTGTTCGGAGCCCTTGAAGGGGTTGTAAGGTTTATCACTGTGCTCACCAGGA 1800  
Qy 1801 GCAAAAGGAAGGTTCTGTCAGAGACGGGCCCAGAGACTAAAGCTTCACTATAATGTGA 1860  
Db 1801 GCAAAAGGAAGGTTCTGTCAGAGACGGGCCCAGAGACTAAAGCTTCACTATAATGTGA 1860  
Qy 1861 CTACAACTAGATTAGTGAATCCAGATCAACTGTTGAAATGAATATGTTCTATCTTCCAA 1920  
Db 1861 CTACAACTAGATTAGTGAATCCAGATCAACTGTTGAAATGAATATGTTCTATCTTCCAA 1920  
Qy 1921 AGAAGCATGTTGACTTTGTACAAAAGGGTTATACACTGAAGTGTCAAAACACAGTCAGATT 1980  
Db 1921 AGAAGCATGTTGACTTTGTACAAAAGGGTTATACACTGAAGTGTCAAAACACAGTCAGATT 1980  
Qy 1981 TTTGGGAAAGTGAACAATTTGAAATTTAGAAAGTGTGCCAGCTTCAAAAACCCGATGTGG 2040  
Db 1981 TTTGGGAAAGTGAACAATTTGAAATTTAGAAAGTGTGCCAGCTTCAAAAACCCGATGTGG 2040  
Qy 2041 TGGGTATCAGGAGGAGCGGCTTAAGGGCGATGCTGGGTTTACAAAAGATTTAGTGGAG 2100  
Db 2041 TGGGTATCAGGAGGAGCGGCTTAAGGGCGATGCTGGGTTTACAAAAGATTTAGTGGAG 2100  
Qy 2101 ACATCCTATCTAGCTGACAGGTATTAATGATGATTTCTTCCATCTTCCCGGATGAGTGTG 2160  
Db 2101 ACATCCTATCTAGCTGACAGGTATTAATGATGATTTCTTCCATCTTCCCGGATGAGTGTG 2160  
Qy 2161 GGTGTGATACAGCTACATAAAAGACTGTTTATGATCGCTTTGATTTTAAAGTTCATTGGAA 2220  
Db 2161 GGTGTGATACAGCTACATAAAAGACTGTTTATGATCGCTTTGATTTTAAAGTTCATTGGAA 2220  
Qy 2221 CTACCAACTTTGTTCTTAAAGAGCTATCTTAAAGCCAAATATCTCTTTGTTTAAACAAA 2280  
Db 2221 CTACCAACTTTGTTCTTAAAGAGCTATCTTAAAGCCAAATATCTCTTTGTTTAAACAAA 2280  
Qy 2281 GATATTTATTTGTTATGAATCTAAATCAAGCCCATCTGCTCAATATGTTACTGTCTTTT 2340  
Db 2281 GATATTTATTTGTTATGAATCTAAATCAAGCCCATCTGCTCAATATGTTACTGTCTTTT 2340

QY 2341 TAATCATGCTGTTTGTATATTAATAATGTTGACCTTCTTAGATTCACCTCCATATG 2400  
DB 2341 TAATCATGCTGTTTGTATATTAATAATGTTGACCTTCTTAGATTCACCTCCATATG 2400  
QY 2401 AATGTAAGCTCTTAACATATGCTCTCTTGTAAATGTAATTTCTTGAATAAACCAT 2460  
DB 2401 AATGTAAGCTCTTAACATATGCTCTCTTGTAAATGTAATTTCTTGAATAAACCAT 2460  
QY 2461 TTGTGAATAT 2470  
DB 2461 TTGTGAATAT 2470

RESULT 13  
ADU05979  
ID ADU05979 standard; DNA; 2470 BP.  
XX AC ADU05979;  
XX AC ADU05979;  
DT 27-JAN-2005 (first entry)  
XX Novel bronchial cancer-associated human gene SegID201.  
DE DE  
XX  
XX KW bronchial cancer; cytostatic; tumour-associated protein;  
KW cancer detection; metastasis; tumour; gene; ds; human.  
XX OS Homo sapiens.  
XX DE10316701-A1.  
PN 04-NOV-2004.  
XX  
PF 09-APR-2003; 2003DB-01016701.  
XX  
PR 09-APR-2003; 2003DB-01016701.  
XX  
PA (HINZ/) HINZMANN B.  
PA (HERM/) HERMANN K.  
PA (CAST/) HEIDEN CASTANOS-VELEZ E.  
XX  
PI Mennerich D, Bruemendorf T, Heiden E, Hermann K, Kinnemann H;  
PI Li X, Roepcke S, Staub E, Hinzmann B, Rosenthal A, Pilarsky C;  
XX  
DR WPI; 2004-786403/78.  
DR P-PSDB; ADU06466.  
XX  
PT New nucleic acid, and derived proteins, useful for diagnosis of bronchial  
PT cancer and in screening for therapeutic and diagnostic agents.  
XX  
PS Claim 1; SEQ ID NO 201; 1381pp; German.  
XX  
CC This invention relates to a novel isolated nucleic acid associated with  
CC bronchial cancer comprising 489 defined sequences given in the  
CC specification. The invention may be useful for the production of  
CC compounds with a cytostatic activity through the inhibition of expression  
CC or activity of tumour-associated proteins. The novel DNA sequences and  
CC the proteins/peptides encoded by them are used for detecting bronchial  
CC cancer or determining the risk of developing it and to screen for  
CC specific binding partners of the DNA or protein sequences, where the  
CC binding partners are potentially useful as agents for treating or  
CC diagnosing bronchial cancer. The DNA or protein sequences can also be  
CC used for prognosis, detection of metastases and for secondary treatment  
CC (of tumours that have been stabilised or are no longer detectable).  
CC Detecting abnormal expression of the DNA sequences provides early  
CC diagnosis of bronchial cancers. The present sequence is that of a novel  
CC bronchial cancer-associated human gene sequence of the invention.  
XX  
SQ Sequence 2470 BP; 767 A; 488 C; 536 G; 679 T; 0 U; 0 Other;  
Query Match 100.0%; Score 2470; DB 13; Length 2470;  
Best Local Similarity 100.0%; Pred No. 0;  
Matches 2470; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTGGCGGGCGGGAAGCGGCCCAACCGCGCATCGAAAGATTCTTAGGAACGCGGTACCA 60  
DB 1 TTGGCGGGCGGGAAGCGGCCCAACCGCGCATCGAAAGATTCTTAGGAACGCGGTACCA 60  
QY 61 GCGCGCTCTCTCAGGACACGAGCGCCCTCTCTTCTGTCGGGCGCGCTCAGCGGTGCC 120  
DB 61 GCGCGCTCTCTCAGGACACGAGCGCCCTCTCTTCTGTCGGGCGCGCTCAGCGGTGCC 120  
QY 121 TCGCGCCCTCAGGTTCTTTTCTTAATTCAAATAAATTCTGCAAGAGACTATGAAAGATT 180  
DB 121 TCGCGCCCTCAGGTTCTTTTCTTAATTCAAATAAATTCTGCAAGAGACTATGAAAGATT 180  
QY 181 ATGATGAACCTTCTCAAAATATTATGAATTAATGAAATTAATGGAACAGGTGCTTGC 240  
DB 181 ATGATGAACCTTCTCAAAATATTATGAATTAATGAAATTAATGGAACAGGTGCTTGC 240  
QY 241 AGGTCAAACTTGCCTGCCATATCTTACCTGGAGAGATGCTAGCTATAAAATCATGATA 300  
DB 241 AGGTCAAACTTGCCTGCCATATCTTACCTGGAGAGATGCTAGCTATAAAATCATGATA 300  
QY 301 AAAACACACTAGGAGTGAATTTGCCCGGATCAAAACCGAGATTGAGGCTTTGAAGAACC 360  
DB 301 AAAACACACTAGGAGTGAATTTGCCCGGATCAAAACCGAGATTGAGGCTTTGAAGAACC 360  
QY 361 TGAGACATCAGCATATATGCTCACTTACCTAGCTGTGTAGAGACAGCAACAAATATTC 420  
DB 361 TGAGACATCAGCATATATGCTCACTTACCTAGCTGTGTAGAGACAGCAACAAATATTC 420  
QY 421 TGGTTCTTGAGTACTGCCCTGGAGAGAGCTGTTTGTGACTATATATTTCCAGGATCGCC 480  
DB 421 TGGTTCTTGAGTACTGCCCTGGAGAGAGCTGTTTGTGACTATATATTTCCAGGATCGCC 480  
QY 481 TGTCAAGAGAGAGACCCGGGTTGTCTTCGGTCAGATAGTATCTGCTGTTGCTTATGTC 540  
DB 481 TGTCAAGAGAGAGACCCGGGTTGTCTTCGGTCAGATAGTATCTGCTGTTGCTTATGTC 540  
QY 541 ACAGCAGGCTATGCTCAAGGACCTCAAGCCAGAAAATTTGCTGTTGATGAATATC 600  
DB 541 ACAGCAGGCTATGCTCAAGGACCTCAAGCCAGAAAATTTGCTGTTGATGAATATC 600  
QY 601 ATAAATTAAGCTGATGACTTTGCTCTGTCGCAAAACCCCAAGGTTACAGGATATAC 660  
DB 601 ATAAATTAAGCTGATGACTTTGCTCTGTCGCAAAACCCCAAGGTTACAGGATATAC 660  
QY 661 ATCTACAGACATGCTGTGGAGTCTGGGCTTATGTCAGCACCTGAGTTAATACAGCAAT 720  
DB 661 ATCTACAGACATGCTGTGGAGTCTGGGCTTATGTCAGCACCTGAGTTAATACAGCAAT 720  
QY 721 CATATCTTGGATCAGAGCAGATGTTGGAGCATGCGGATCTGTTATATGTTCTTATGT 780  
DB 721 CATATCTTGGATCAGAGCAGATGTTGGAGCATGCGGATCTGTTATATGTTCTTATGT 780  
QY 781 GTGGATTTCTACCATTTTCATGATGATGATGATGATGATGATGATGATGATGATGATG 840  
DB 781 GTGGATTTCTACCATTTTCATGATGATGATGATGATGATGATGATGATGATGATGATG 840  
QY 841 GAAATATGATGTTCCCAAGTGGCTCTCTCCAGTAGCATTTCTGCTTCTTCAACAAATGC 900  
DB 841 GAAATATGATGTTCCCAAGTGGCTCTCTCCAGTAGCATTTCTGCTTCTTCAACAAATGC 900  
QY 901 TGCAGGTGACCCAAAGAAACCGATTTCTATGAAATATCTATGAAACCATCTCGATCA 960  
DB 901 TGCAGGTGACCCAAAGAAACCGATTTCTATGAAATATCTATGAAACCATCTCGATCA 960  
QY 961 TGCAGATTAACATCTATCTCTGTTGAGTGCAGCAAGCAATCTTTTATTCACCTCGATG 1020  
DB 961 TGCAGATTAACATCTATCTCTGTTGAGTGCAGCAAGCAATCTTTTATTCACCTCGATG 1020  
QY 1021 ATGATTGGCTTAACGAACCTTTCTGTACATCAAGAAACCAAGGCAACCAATGAGGATT 1080  
DB 1021 ATGATTGGCTTAACGAACCTTTCTGTACATCAAGAAACCAAGGCAACCAATGAGGATT 1080  
QY 1081 TAATTTCACTGTGGCAGTATGATCACTACGGCTACCTATCTTCTTCTTAGCCAAGA 1140

Db 1081 |||||TAATTTCACTGTGGCAGTAGATGATCACTCAGCGCTACCTATCTTCTGCTTAGCCAGA 1140  
Qy 1141 AGGCTCGGGGAAACACAGTTCGTTTAAAGCTTCTTCTTCTCTGTCGACAGCCAGTG 1200  
Db 1141 AGGCTCGGGGAAACACAGTTCGTTTAAAGCTTCTTCTTCTCTGTCGACAGCCAGTG 1200  
Qy 1201 CTACCCCATTTACAGACATCAAGTCAAAATAATTGGAGTCTGGAAGATGTGACCGCAAGTG 1260  
Db 1201 CTACCCCATTTACAGACATCAAGTCAAAATAATTGGAGTCTGGAAGATGTGACCGCAAGTG 1260  
Qy 1261 ATAAAAATATGTGGGGGAAATTAATAGACTATGATTTGGTGTGAAGATGAATTTATCAACAG 1320  
Db 1261 ATAAAAATATGTGGGGGAAATTAATAGACTATGATTTGGTGTGAAGATGAATTTATCAACAG 1320  
Qy 1321 GTGCTGCTACTCCCGAAACATCAAGTTCACAACTACCAAGTCTGGAAGATGAATTTATCAACAG 1380  
Db 1321 GTGCTGCTACTCCCGAAACATCAAGTTCACCAAGTCTGGAAGATGAATTTATCAACAG 1380  
Qy 1381 AATCTAAATCATTAATCTCCAGCCTTATGCAAGACACTGCAAAATAATTAAAGAACAAAG 1440  
Db 1381 AATCTAAATCATTAATCTCCAGCCTTATGCAAGACACTGCAAAATAATTAAAGAACAAAG 1440  
Qy 1441 AAAATGTATATCTCTTAAGTCTGCTGTAAAGATGAAGTACTTTATGTTTCTGAGC 1500  
Db 1441 AAAATGTATATCTCTTAAGTCTGCTGTAAAGATGAAGTACTTTATGTTTCTGAGC 1500  
Qy 1501 CAAAGACTCCAGTTAATGAAGACAGCATGAAGAGAGAAATCTACTACGCCAAATCGTT 1560  
Db 1501 CAAAGACTCCAGTTAATGAAGACAGCATGAAGAGAGAAATCTACTACGCCAAATCGTT 1560  
Qy 1561 ACATACACCTCAAAAGCTAGAAACAGTCCGTGAAAGAACTCCAAATTAATAATACCAG 1620  
Db 1561 ACATACACCTCAAAAGCTAGAAACAGTCCGTGAAAGAACTCCAAATTAATAATACCAG 1620  
Qy 1621 TAAATTCACAGGAACAGCAAGTTAATGA CAGGTGTCTATTAGCCCTGAGAGCGGTGCC 1680  
Db 1621 TAAATTCACAGGAACAGCAAGTTAATGA CAGGTGTCTATTAGCCCTGAGAGCGGTGCC 1680  
Qy 1681 GCTCAGTGAATTTGATCTCAACAGCAGCATATGAGAGAGACTCCAAAAGAAAGGGAG 1740  
Db 1681 GCTCAGTGAATTTGATCTCAACAGCAGCATATGAGAGAGACTCCAAAAGAAAGGGAG 1740  
Qy 1741 CCAAAGTGTGGGAGCCTTGAAGGGGTTGATAGAGTTATCAGTGTCTCACCAGGA 1800  
Db 1741 CCAAAGTGTGGGAGCCTTGAAGGGGTTGATAGAGTTATCAGTGTCTCACCAGGA 1800  
Qy 1801 GCAAAAGGAAGGTTCTGCCAGAGACGGGCCCAGAGACTTAAAGCTTCACTATAATGTGA 1860  
Db 1801 GCAAAAGGAAGGTTCTGCCAGAGACGGGCCCAGAGACTTAAAGCTTCACTATAATGTGA 1860  
Qy 1861 CTACAACTAGATTAGTGAATCCAGATCAACTGTTGAATGAATGAATGCTATTCTTCCAA 1920  
Db 1861 CTACAACTAGATTAGTGAATCCAGATCAACTGTTGAATGAATGAATGCTATTCTTCCAA 1920  
Qy 1921 AGAAGCATGTTGACTTTGTACAAAAGGTTATACACTGAAGTGTCAAAACAGTCAAGATT 1980  
Db 1921 AGAAGCATGTTGACTTTGTACAAAAGGTTATACACTGAAGTGTCAAAACAGTCAAGATT 1980  
Qy 1981 TTGGGAAAGTGACAATGCAATTTGAAATTAAGAGTGTGCCAGCTTCAAAAACCCGATGTGG 2040  
Db 1981 TTGGGAAAGTGACAATGCAATTTGAAATTAAGAGTGTGCCAGCTTCAAAAACCCGATGTGG 2040  
Qy 2041 TGGGTATCAGAGGACGCGCTTAAGGCCATCGCTGGGTTTACAAAAGTTAGTGGAG 2100  
Db 2041 TGGGTATCAGAGGACGCGCTTAAGGCCATCGCTGGGTTTACAAAAGTTAGTGGAG 2100  
Qy 2101 ACATCTATCTAGCTGCAAGGTATATTTAGTGAATCTTCCATCTCGCGGATGAGTGTG 2160  
Db 2101 ACATCTATCTAGCTGCAAGGTATATTTAGTGAATCTTCCATCTCGCGGATGAGTGTG 2160  
Qy 2161 GGTGTGATACAGCCTACATAAAGACTGTTTATGATCGCTTTGATTTTAAAGTTCATTGGAA 2220

Db 2161 GGTGTGATACAGCCTACATAAAGACTGTTATGATCGCTTTGATTTTAAAGTTCATTGGAA 2220  
Qy 2221 CTACCAACTGTTTCTAAAGAGCTATCTTAAGACCAATATCTTTTGTGTTTAAACAAA 2280  
Db 2221 CTACCAACTGTTTCTAAAGAGCTATCTTAAGACCAATATCTTTTGTGTTTAAACAAA 2280  
Qy 2281 GATATTATTTGTGTATGAATCTAAATCAAGCCCATCTGTCAATTAATGTTACTGTTT 2340  
Db 2281 GATATTATTTGTGTATGAATCTAAATCAAGCCCATCTGTCAATTAATGTTACTGTTT 2340  
Qy 2341 TAATCATGTGGTTTGTATATAATTAATTAATTTGTTGACTTCTTAGATTCATTCATATGTG 2400  
Db 2341 TAATCATGTGGTTTGTATATAATTAATTAATTTGTTGACTTCTTAGATTCATTCATATGTG 2400  
Qy 2401 AATGTAAGCTCTTAAGTATGCTCTTTGTAATGTTGTAATTTCTTCTGAAATAAACCAT 2460  
Db 2401 AATGTAAGCTCTTAAGTATGCTCTTTGTAATGTTGTAATTTCTTCTGAAATAAACCAT 2460  
Qy 2461 TTGTGAATAT 2470  
Db 2461 TTGTGAATAT 2470  
RESULT 14  
ADX07296  
ID ADX07296 standard; DNA; 2470 BP.  
XX ADX07296;  
AC ADX07296;  
XX 21-APR-2005 (first entry)  
DT Cyclin-dependent kinase modulation biomarker DNA SEQ ID NO 1861.  
XX Cytostatic; cyclin-dependent kinase; cdk; biomarker; gene; ds.  
XX Homo sapiens.  
OS WO2005012875-A2.  
XX 10-FEB-2005.  
PD 29-JUL-2004; 2004WO-US024424.  
PF 29-JUL-2003; 2003US-0490890P.  
PR (BRIM ) BRISTOL-MYERS SQUIBB CO.  
PA Li M, Rupnow BA, Webster KR, Jackson DG, Wong TW;  
XX WPI; 2005-163068/17.  
PI P-PSDB; ADX07297.  
PT Biomarkers useful for predicting or determining the response of a mammal to a cancer treatment comprising administration of a modulator of cyclin-dependent kinase activity.  
PS Claim 5; SEQ ID NO 1861; 141pp; English.  
XX This invention describes a novel method of predicting or determining whether a mammal will respond or is responding to an anti-cancer agent that modulates cyclin-dependent kinase (cdk) activity. The method comprises measuring the level of one or more biomarkers selected from 2774 biomarkers given in the specification (nucleotide sequence SEQ ID NO:1246 (Genbank EST W28729) is especially preferred). The method of the invention is utilized in a kit for determining or predicting whether a patient would be susceptible or resistant to treatment by an agent modulating cdk activity. The invention also describes a method for utilizing individualized genetic profiles for treating diseases and disorders based on patient's response and molecular level, specialized microarrays comprising the biomarkers described, antibodies directed against the biomarkers and a cell culture model to identify biomarkers. The cdk modulator is preferably N-5-[(5-(1,1-Dimethylethyl)-2-oxazoly]methylthio]-2-thiazolyl-4-piperidine carboxamide, 0.5-L-

CC tartaric acid salt. Note: The sequence data for this patent did not form  
CC part of the printed specification, but was obtained in electronic format  
CC directly from WIPO at ftp.wipo.int/pub/published pct sequences. This  
CC sequence encodes a biomarker used in the method of the invention.

XX  
SQ Sequence 2470 BP; 767 A; 488 C; 536 G; 679 T; 0 U; 0 Other;

Query Match 100.0%; Score 2470; DB 14; Length 2470;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2470; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 TTGGCGGGGGAAGCGGCCACACACCGCGGATCGAAAGATTCTTAGGAACGCGTACCA 60
DB 1 TTGGCGGGGGAAGCGGCCACACACCGCGGATCGAAAGATTCTTAGGAACGCGGTACCA 60
QY 61 GCGCGGTCTCTCAGGACAGCAGCGCCCTGTCTTCTGTGGGCGCGCTCAGCGGTGCC 120
DB 61 GCGCGGTCTCTCAGGACAGCAGCGCCCTGTCTTCTGTGGGCGCGCTCAGCGGTGCC 120
QY 121 TCGGCGCCCTCAGGTTCTTTTCTTAATTCMAATAAACTTGCAGAGGACTATGAAGAATT 180
DB 121 TCGGCGCCCTCAGGTTCTTTTCTTAATTCMAATAAACTTGCAGAGGACTATGAAGAATT 180
QY 181 ATGATCAACTTCTCAAAATATTATGAATTACATGAACATTTGGGACAGTGGCTTGCAA 240
DB 181 ATGATCAACTTCTCAAAATATTATGAATTACATGAACATTTGGGACAGTGGCTTGCAA 240
QY 241 AGGTCAAACTTGGCTGCCATATCTTACTTGGAGAGATGGTAGTATAAAAAATCATGGATA 300
DB 241 AGGTCAAACTTGGCTGCCATATCTTACTTGGAGAGATGGTAGTATAAAAAATCATGGATA 300
QY 301 AAAACACATAGGAGTGTATTTGCCCGGATCAAAACGGAGATTGAGGCTTTGAAGAACC 360
DB 301 AAAACACATAGGAGTGTATTTGCCCGGATCAAAACGGAGATTGAGGCTTTGAAGAACC 360
QY 361 TGAGACATCAGCATATATGTCAACTCTACATGTGTAGAGACAGCAACAAATATTCA 420
DB 361 TGAGACATCAGCATATATGTCAACTCTACATGTGTAGAGACAGCAACAAATATTCA 420
QY 421 TGGTTCTTGAGTACTGCCCTGGAGGAGAGCTGTTTGCATATATAATTTCCAGGATCGCC 480
DB 421 TGGTTCTTGAGTACTGCCCTGGAGGAGAGCTGTTTGCATATATAATTTCCAGGATCGCC 480
QY 481 TGTCAAGAGAGAGACCGGGTGTCTTCGGTCAGATAGTATCTGCTGTGTTGCTTATGTGC 540
DB 481 TGTCAAGAGAGAGACCGGGTGTCTTCGGTCAGATAGTATCTGCTGTGTTGCTTATGTGC 540
QY 541 ACAGCCAGGGCTATGCTCAGGGACCTCAAGCCAGAAAATTTGCTGTTTGTATGATATC 600
DB 541 ACAGCCAGGGCTATGCTCAGGGACCTCAAGCCAGAAAATTTGCTGTTTGTATGATATC 600
QY 601 ATAAATTAAAGCTGATTGACTTTGGTCTCTGTGCAAAACCCCAAGGTAACAAGGATTACC 660
DB 601 ATAAATTAAAGCTGATTGACTTTGGTCTCTGTGCAAAACCCCAAGGTAACAAGGATTACC 660
QY 661 ATCTACAGACATGCTGTGGAGTCTGGCTTTATGCAGCACCTGAGTTAAATACAAGGCAAT 720
DB 661 ATCTACAGACATGCTGTGGAGTCTGGCTTTATGCAGCACCTGAGTTAAATACAAGGCAAT 720
QY 721 CATATCTTGGATCAGAGGCAGATGTTTGGAGCATGCGCATCTGTTATATGTTCTTATGT 780
DB 721 CATATCTTGGATCAGAGGCAGATGTTTGGAGCATGCGCATCTGTTATATGTTCTTATGT 780
QY 781 GTGGATTTCTACATTTGATGATGATTAATGTAATGGCTTTTATACAAGAAGATTATGAGAG 840
DB 781 GTGGATTTCTACATTTGATGATGATTAATGTAATGGCTTTTATACAAGAAGATTATGAGAG 840
QY 841 GAAAAATATGATGTTCCCAAGTGGCTCTCTCCAGTAGCATCTTCTTCTTCAACAAATGC 900
DB 841 GAAAAATATGATGTTCCCAAGTGGCTCTCTCCAGTAGCATCTTCTTCTTCAACAAATGC 900
QY 901 TCGAGGTGACCAAGAAACCGGATTTCTATGAAAAATCTATGAAACCATCTCCCTGGATCA 960
DB 901 TCGAGGTGACCAAGAAACCGGATTTCTATGAAAAATCTATGAAACCATCTCCCTGGATCA 960
```

```
DB 901 TCGAGGTGACCAAGAAACCGGATTTCTATGAAAAATCTATGAAACCATCTCCCTGGATCA 960
QY 961 TCGAAGATTATCAACTATCTCTGTGTAGTGCAAGCAAGAAATCTTTTATTTACCTCGATG 1020
DB 961 TCGAAGATTATCAACTATCTCTGTGTAGTGCAAGCAAGAAATCTTTTATTTACCTCGATG 1020
QY 1021 ATGATTTGCGTAAACAGAACCTTTCTGTATCATCAGAGAAACAAACAGGCAAAATATGGAGATT 1080
DB 1021 ATGATTTGCGTAAACAGAACCTTTCTGTATCATCAGAGAAACAAACAGGCAAAATATGGAGATT 1080
QY 1081 TAAATTTCACTGTGGCAGTATGATCACTCAGCGCTACCTTCTCTGCTTCTTAGCCAAAGA 1140
DB 1081 TAAATTTCACTGTGGCAGTATGATCACTCAGCGCTACCTTCTCTGCTTCTTAGCCAAAGA 1140
QY 1141 AGGCTCGGGGAAAAACAGTTCGTTTAAAGGCTTTCTTTCTCTCTGTGGAAAGCCAGTGG 1200
DB 1141 AGGCTCGGGGAAAAACAGTTCGTTTAAAGGCTTTCTTTCTCTCTGTGGAAAGCCAGTGG 1200
QY 1201 CTACCCCATTTACAGACATCAAGTCAAAATAATTTGGAGTCTGGAAGNTGAGCCGCAAGTG 1260
DB 1201 CTACCCCATTTACAGACATCAAGTCAAAATAATTTGGAGTCTGGAAGNTGAGCCGCAAGTG 1260
QY 1261 ATAAAAAATTTATCTGCGGGGATTAATAGACTATGATTTGGTGTGAAGATATTTTATCAACAG 1320
DB 1261 ATAAAAAATTTATCTGCGGGGATTAATAGACTATGATTTGGTGTGAAGATATTTTATCAACAG 1320
QY 1321 GTGCTGCTACTCCCGAACATCACAGTTTACCAAGTACTTGGACAGAAATCAAAATGGGGTGG 1380
DB 1321 GTGCTGCTACTCCCGAACATCACAGTTTACCAAGTACTTGGACAGAAATCAAAATGGGGTGG 1380
QY 1381 AATCTAAATCATTAATCTCAGGCTTATGCAAGAACCTTGCACAAATAAATTTAAAGAACAAAG 1440
DB 1381 AATCTAAATCATTAATCTCAGGCTTATGCAAGAACCTTGCACAAATAAATTTAAAGAACAAAG 1440
QY 1441 AAAATGTAATACTCCTTAAGTCTGTAAAGAAATGAAGAGTACTTATGTTTCTCTGAGC 1500
DB 1441 AAAATGTAATACTCCTTAAGTCTGTAAAGAAATGAAGAGTACTTATGTTTCTCTGAGC 1500
QY 1501 CAAAGACTCCAGTTAATAAGAACCCAGCATATGAGAGAGAAATATCTACTACGCCAAATCGTT 1560
DB 1501 CAAAGACTCCAGTTAATAAGAACCCAGCATATGAGAGAGAAATATCTACTACGCCAAATCGTT 1560
QY 1561 ACATCTACACCTTCAAAAGCTAGAAAACAGTGCCTGAAAGAAAATCTCAATTTAAATATCCAG 1620
DB 1561 ACATCTACACCTTCAAAAGCTAGAAAACAGTGCCTGAAAGAAAATCTCAATTTAAATATCCAG 1620
QY 1621 TAAATTTCAACAGGAAACAGACAAAGTTAATGACAGGTGTCAATAGCCCTGAGAGCCGTGCC 1680
DB 1621 TAAATTTCAACAGGAAACAGACAAAGTTAATGACAGGTGTCAATAGCCCTGAGAGCCGTGCC 1680
QY 1681 GCTCAGTGAATTTGGATCTCAACCAAGCACATATGAGAGAGACTCCAAAAGAAAAGGGAG 1740
DB 1681 GCTCAGTGAATTTGGATCTCAACCAAGCACATATGAGAGAGACTCCAAAAGAAAAGGGAG 1740
QY 1741 CCAAGAGTGTGTGGAGCCTTGAAGAGGGGTGGATAAGGTTATCACTGTGCTCACACAGA 1800
DB 1741 CCAAGAGTGTGTGGAGCCTTGAAGAGGGGTGGATAAGGTTATCACTGTGCTCACACAGA 1800
QY 1801 GCAAAAGGAGGGTTCTGCCAGAGACGGGCCAGAGACTAAAGCTTCACTATAATGTGA 1860
DB 1801 GCAAAAGGAGGGTTCTGCCAGAGACGGGCCAGAGACTAAAGCTTCACTATAATGTGA 1860
QY 1861 CTACAACTAGATTAGTGAATCCAGATCAACTGTTGAATGAATAATATGTTCTTCCAA 1920
DB 1861 CTACAACTAGATTAGTGAATCCAGATCAACTGTTGAATGAATAATATGTTCTTCCAA 1920
QY 1921 AGAAGCATGTTGACTTTGTACAAAAGGGTTATACACTGAAAGTGTCAACACAGTCAGATT 1980
DB 1921 AGAAGCATGTTGACTTTGTACAAAAGGGTTATACACTGAAAGTGTCAACACAGTCAGATT 1980
QY 1981 TTGGGAAAAGTGAACAATTTGAAATTTAGAAAGTGTGCGAGCTTCAAAAACCCGATGTGG 2040
DB 1981 TTGGGAAAAGTGAACAATTTGAAATTTAGAAAGTGTGCGAGCTTCAAAAACCCGATGTGG 2040
```

Qy	2041	TGGGTATCAGGAGCAGCGGCTTAAGGGCGATGCTCGGGTTTACAAAGATTAGTGAAG	2100
Db	2041	TGGGTATCAGGAGCAGCGGCTTAAGGGCGATGCTCGGGTTTACAAAGATTAGTGAAG	2100
Qy	2101	ACATCCTATCTAGCTGCAAGGTATAATTGATGGATTCTTCATCCTGCCGGATGAGTGTG	2160
Db	2101	ACATCCTATCTAGCTGCAAGGTATAATTGATGGATTCTTCATCCTGCCGGATGAGTGTG	2160
Qy	2161	GGTGTGATCAGCCCTACATAAAGACTGTATTGATGCGCTTTTGATTTAAAGTTCATTGGAA	2220
Db	2161	GGTGTGATCAGCCCTACATAAAGACTGTATTGATGCGCTTTTGATTTAAAGTTCATTGGAA	2220
Qy	2221	CTACCAACTTGTGTTTCTAAGAGCTATCTTAAGACCAATATCTCTTGTGTTTTTAAACAAAA	2280
Db	2221	CTACCAACTTGTGTTTCTAAGAGCTATCTTAAGACCAATATCTCTTGTGTTTTTAAACAAAA	2280
Qy	2281	GATATTAATTTTGATGTAATCTAAATCAAGCCCATCTGTCATATGTTACTGTCCTTTT	2340
Db	2281	GATATTAATTTTGATGTAATCTAAATCAAGCCCATCTGTCATATGTTACTGTCCTTTT	2340
Qy	2341	TAATCATGTGTTTTTGATATTAATAATGTTGACTTTCTTAGAATTCACATTCACATATGTG	2400
Db	2341	TAATCATGTGTTTTTGATATTAATAATGTTGACTTTCTTAGAATTCACATTCACATATGTG	2400
Qy	2401	AATGTAAGCTCTTAACTATGTCCTTTTGTAATGTGTAATTTCTTTCTGAAATAAAACCAT	2460
Db	2401	AATGTAAGCTCTTAACTATGTCCTTTTGTAATGTGTAATTTCTTTCTGAAATAAAACCAT	2460
Qy	2461	TTGTGAATAT 2470	
Db	2461	TTGTGAATAT 2470	

RESULT 15  
ADW93565

ADW93565	standard; cDNA; 2470 BP.
ADW93565;	
05-MAY-2005	(first entry)
Human maternal embryonic leucine zipper kinase, MELK, cDNA #1.	
Maternal embryonic leucine zipper kinase; ss; gene; drug screening; angiogenesis disorder; antiangiogenic; cardiovascular disease; apoptosis; cancer; cytostatic; neoplasm; diagnosis; gene therapy.	
Homo sapiens.	
Key	Location/Qualifiers
CDS	171..2126
	/*tag= a
	/product= "MELK"
WO2005016279-A2.	
24-FEB-2005.	
12-AUG-2004; 2004WO-US026231.	
14-AUG-2003; 2003US-0495193P.	
(EXEL-) EXELIXIS INC.	
Kadyk L, Francis GR, Heuer TS, Lickteig K;	
WPI; 2005-173218/18.	
P-PSDB; ADW93570.	
Identifying candidate RAC pathway modulating agents useful for diagnosing or treating e.g. cancer, comprises screening for agents that modulate the activity of Maternal Embryonic Leucine Zipper Kinase (MELK).	





This Page Blank (uspto)

GenCore version 5.1.9  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 21, 2006, 19:38:50 ; Search time 458 Seconds  
(without alignments)  
10090.914 Million cell updates/sec

Title: US-10-656-598-1

Perfect score: 2470

Sequence: 1 t699cgggaggaagggccca.....ataaacatttgtgaatat 2470

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 1403666 seqs, 935554401 residues

Total number of hits satisfying chosen parameters: 2807332

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA.\*

- 1: /EMC\_Celerra\_SIDS3/ptodata/2/ina/1 COMB.seq.\*
- 2: /EMC\_Celerra\_SIDS3/ptodata/2/ina/5 COMB.seq.\*
- 3: /EMC\_Celerra\_SIDS3/ptodata/2/ina/6A COMB.seq.\*
- 4: /EMC\_Celerra\_SIDS3/ptodata/2/ina/6B COMB.seq.\*
- 5: /EMC\_Celerra\_SIDS3/ptodata/2/ina/7 COMB.seq.\*
- 6: /EMC\_Celerra\_SIDS3/ptodata/2/ina/H COMB.seq.\*
- 7: /EMC\_Celerra\_SIDS3/ptodata/2/ina/PCTUS COMB.seq.\*
- 8: /EMC\_Celerra\_SIDS3/ptodata/2/ina/PP COMB.seq.\*
- 9: /EMC\_Celerra\_SIDS3/ptodata/2/ina/RE COMB.seq.\*
- 10: /EMC\_Celerra\_SIDS3/ptodata/2/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2470	100.0	2470	4	US-09-880-107-1725 Sequence 1725, Ap
2	2470	100.0	2470	4	US-09-870-937-9 Sequence 9, Appli
3	214	8.7	245	3	US-09-513-999C-11495 Sequence 11495, A
4	206.8	8.4	1863	3	US-09-949-016-4303 Sequence 4303, Ap
5	197.6	8.0	2224	3	US-09-949-016-2384 Sequence 2384, Ap
6	197.6	8.0	2224	3	US-09-949-016-2385 Sequence 2385, Ap
7	197.6	8.0	2950	3	US-09-949-016-1546 Sequence 1546, Ap
8	197.6	8.0	2950	3	US-09-949-016-1547 Sequence 1547, Ap
9	196	7.9	2175	3	US-09-984-890-1 Sequence 1, Appli
10	196	7.9	2175	3	US-10-274-194-1 Sequence 1, Appli
11	196	7.9	2175	3	US-10-760-407-1 Sequence 1, Appli
12	188.8	7.6	1647	3	US-09-101-146-44 Sequence 44, Appl
13	188.6	7.6	1736	3	US-08-557-006C-37 Sequence 37, Appl
14	188.6	7.6	1783	3	US-08-557-006C-36 Sequence 36, Appl
15	181.2	7.3	2210	3	US-09-949-016-1699 Sequence 1699, Ap
16	173.8	7.0	2698	2	US-08-677-298-1 Sequence 1, Appli
17	173.8	7.0	2912	3	US-09-949-016-2363 Sequence 2363, Ap
18	172.6	7.0	1742	3	US-08-557-006C-38 Sequence 38, Appl
19	172.6	7.0	2652	3	US-08-557-006C-39 Sequence 39, Appl
20	172.6	7.0	2761	3	US-08-557-006C-24 Sequence 24, Appl
21	170.6	6.9	2914	3	US-09-949-016-343 Sequence 343, App
22	170.4	6.9	3609	3	US-09-799-875-6 Sequence 6, Appli
23	170.4	6.9	5983	3	US-09-799-875-4 Sequence 4, Appli

24	167.8	6.8	1539	3	US-09-633-328B-1 Sequence 1, Appli
25	159.4	6.5	3857	3	US-10-464-939-1 Sequence 1, Appli
26	156	6.3	1929	3	US-09-359-161-4 Sequence 4, Appli
27	155.4	6.3	4790	3	US-10-464-939-3 Sequence 3, Appli
28	149.8	6.1	1590	3	US-09-554-726A-4 Sequence 4, Appli
29	149.8	6.1	2257	3	US-09-554-726A-6 Sequence 6, Appli
30	149.6	6.1	2738	3	US-09-554-726A-9 Sequence 9, Appli
31	149.2	6.0	1254	3	US-09-248-796A-4338 Sequence 4338, Ap
32	149.2	6.0	1747	3	US-08-557-006C-44 Sequence 44, Appl
33	148	6.0	1787	3	US-09-554-726A-13 Sequence 13, Appl
34	147	6.0	2334	3	US-10-003-690-3 Sequence 3, Appli
35	147	6.0	2337	3	US-10-116-326-1 Sequence 1, Appli
36	147	6.0	2337	3	US-10-803-277-1 Sequence 1, Appli
37	147	6.0	2980	3	US-10-003-690-1 Sequence 1, Appli
38	145	5.9	1902	3	US-09-487-558B-337 Sequence 337, App
39	143.2	5.8	2899	3	US-09-774-528-112 Sequence 112, App
40	143.2	5.8	2899	3	US-10-120-988-112 Sequence 112, App
41	139.8	5.7	2112	3	US-10-116-326-5 Sequence 5, Appli
42	139.8	5.7	2112	3	US-10-803-277-5 Sequence 5, Appli
43	137.6	5.6	2902	3	US-09-579-664B-4 Sequence 4, Appli
44	137.6	5.6	2902	3	US-10-355-975A-4 Sequence 4, Appli
45	137.6	5.6	2902	5	US-09-980-464-4 Sequence 4, Appli

ALIGNMENTS

RESULT 1

US-09-880-107-1725  
; Sequence 1725, Application US/09880107  
; Patent No. 6974667  
; GENERAL INFORMATION:  
; APPLICANT: Horne, Darci T.  
; APPLICANT: Vockley, Joseph G.  
; APPLICANT: Scharf, Uwe  
; APPLICANT: Gene Logic, Inc.  
; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer  
; FILE REFERENCE: 44921-5028-WO  
; CURRENT APPLICATION NUMBER: US/09/880,107  
; CURRENT FILING DATE: 2001-06-14  
; PRIOR APPLICATION NUMBER: US 60/211,379  
; PRIOR FILING DATE: 2000-06-14  
; PRIOR APPLICATION NUMBER: US 60/237,054  
; PRIOR FILING DATE: 2000-10-02  
; NUMBER OF SEQ ID NOS: 3950  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 1725  
; LENGTH: 2470  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: Genbank Accession No. 6974667 D79997  
US-09-880-107-1725

Query Match 100.0%; Score 2470; DB 4; Length 2470;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2470; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	TTGGCGGGGGAAGCGGCACACCGGGCGATCGAAGATTCTTAGGAACCGCGTACCA	60
Db	1	TTGGCGGGGGAAGCGGCACACCGGGCGATCGAAGATTCTTAGGAACCGCGTACCA	60
Qy	61	GC CGCGTCTCTCAGACACAGCGCCCTGTCTCTCTGTGGCGCGCGCTCAGCGGTGCC	120
Db	61	GC CGCGTCTCTCAGACACAGCGCCCTGTCTCTCTGTGGCGCGCGCTCAGCGGTGCC	120
Qy	121	TC CGCCCTCAGGTTCTTTTCTAATTCAAATAAATTGCAAGAGGACTATGAAGATT	180
Db	121	TC CGCCCTCAGGTTCTTTTCTAATTCAAATAAATTGCAAGAGGACTATGAAGATT	180
Qy	181	ATGATGAACCTCTCAATATATGATTAATGCAAGACTATTGGGACAGTGGCTTGCA	240
Db	181	ATGATGAACCTCTCAATATATGATTAATGCAAGACTATTGGGACAGTGGCTTGCA	240

QY	241	AGGTCAAACTTGGCTGCCATATCCTTACTTGGAGAGATGGTAGCTATAAAAAATCATGGATA	300
Db	241		
QY	241	AGGTCAAACTTGGCTGCCATATCCTTACTGGAAGATGGTAGCTATAAAAAATCATGGATA	300
Db	241		
QY	301	AAACACACATAGGAGTGATTTCCCGGGATCAAAACGGAGATTTGAGGCGCTTGAAGAAC	360
Db	301		
QY	301	AAACACACATAGGAGTGATTTCCCGGGATCAAAACGGAGATTTGAGGCGCTTGAAGAAC	360
Db	301		
QY	361	TGAGACATCAGCATATATGTCAACTCTACCATGTCTAGAGACAGCCAAACAAATATTCA	420
Db	361		
QY	361	TGAGACATCAGCATATATGTCAACTCTACCATGTCTAGAGACAGCCAAACAAATATTCA	420
Db	361		
QY	421	TGGTCTTGTAGTACTGCCCTGGAGGAGAGCTGTTTGACTATATAAAATTTCCAGGATCGCC	480
Db	421		
QY	421	TGGTCTTGTAGTACTGCCCTGGAGGAGAGCTGTTTGACTATATAAAATTTCCAGGATCGCC	480
Db	421		
QY	481	TGTCAGAAGAGAGACCCGGGTTGTCTTCGGTCAGATAGTATCTGCTGTGCTTATGTGC	540
Db	481		
QY	481	TGTCAGAAGAGAGACCCGGGTTGTCTTCGGTCAGATAGTATCTGCTGTGCTTATGTGC	540
Db	481		
QY	541	ACAGCCAGGGCTATGCTCAGGGACCTCAAGCCAGAAAAATTTGCTGTTGATGAATATC	600
Db	541		
QY	541	ACAGCCAGGGCTATGCTCAGGGACCTCAAGCCAGAAAAATTTGCTGTTGATGAATATC	600
Db	541		
QY	601	ATAAATTTAAAGCTGATTTGACTTTGGTCTCTGTGCAAAACCCAAAGGGTTAAACAAGGATTACC	660
Db	601		
QY	601	ATAAATTTAAAGCTGATTTGACTTTGGTCTCTGTGCAAAACCCAAAGGGTTAAACAAGGATTACC	660
Db	601		
QY	661	ATCTACAGACATGCTGTGGAGTCTGGCTTTATGCAGCACCTGAGTTAATACAAGGCAAAAT	720
Db	661		
QY	661	ATCTACAGACATGCTGTGGAGTCTGGCTTTATGCAGCACCTGAGTTAATACAAGGCAAAAT	720
Db	661		
QY	721	CATATCTTGATCAGAGGCAGATGTTTGGAGCATGGGCATACTGTTTATATGTTCTTATGT	780
Db	721		
QY	721	CATATCTTGATCAGAGGCAGATGTTTGGAGCATGGGCATACTGTTTATATGTTCTTATGT	780
Db	721		
QY	781	GTGGATTTCTACCAATTTGATGATGATAATGTAATGGCTTTATACAAGAAGATTATGAGAG	840
Db	781		
QY	781	GTGGATTTCTACCAATTTGATGATGATAATGTAATGGCTTTATACAAGAAGATTATGAGAG	840
Db	781		
QY	841	GAAAAATATGATGTCCCAAGTGGCTCTCTCCAGTAGCATTCTGCTTCTTCAACAAATGC	900
Db	841		
QY	841	GAAAAATATGATGTCCCAAGTGGCTCTCTCCAGTAGCATTCTGCTTCTTCAACAAATGC	900
Db	841		
QY	901	TGAGGTGGACCCAAAGAACGGAATTTCTATGAAAAATCTTATTTGAACCATCCCTGGATCA	960
Db	901		
QY	901	TGAGGTGGACCCAAAGAACGGAATTTCTATGAAAAATCTTATTTGAACCATCCCTGGATCA	960
Db	901		
QY	961	TGCAAGATTACAATCTATCTGTTGAGTGGCAAGCAAGAAATCCTTTTATTCACCTCGATG	1020
Db	961		
QY	961	TGCAAGATTACAATCTATCTGTTGAGTGGCAAGCAAGAAATCCTTTTATTCACCTCGATG	1020
Db	961		
QY	1021	ATGATTTGCGTAAACAGAACTTTCTGTACATCACAGAAAAACAGGCAAAACAAATGGAGATT	1080
Db	1021		
QY	1021	ATGATTTGCGTAAACAGAACTTTCTGTACATCACAGAAAAACAGGCAAAACAAATGGAGATT	1080
Db	1021		
QY	1081	TAATTTTCACTGTGGCAGTATGATCACCCTCAGGCTACCTATCTTCTGCTTCTAGCCAGA	1140
Db	1081		
QY	1081	TAATTTTCACTGTGGCAGTATGATCACCCTCAGGCTACCTATCTTCTGCTTCTAGCCAGA	1140
Db	1081		
QY	1141	AGGCTGGGGAAAAACAGTTTCTTTTAAAGGCTTTCTTCTTCTCTGTGGACAAGCCAGTG	1200
Db	1141		
QY	1141	AGGCTGGGGAAAAACAGTTTCTTTTAAAGGCTTTCTTCTTCTCTGTGGACAAGCCAGTG	1200
Db	1141		
QY	1201	CTACCCCATTTACAGACATCAAGTCAAAATATTTGGAGTCTGGAAGATGTCAACCGCAAGTG	1260
Db	1201		
QY	1201	CTACCCCATTTACAGACATCAAGTCAAAATATTTGGAGTCTGGAAGATGTCAACCGCAAGTG	1260
Db	1201		
QY	1261	ATAAATTTATGTGGCGGATTAATAGACTATGATTTGGTGTGAAGATGATTTATCAACAG	1320
Db	1261		
QY	1261	ATAAATTTATGTGGCGGATTAATAGACTATGATTTGGTGTGAAGATGATTTATCAACAG	1320
Db	1261		

QY	1321	GTGCTGCTACTCCCGAACATCACAGTTTACCAAGTACTGGAAGATCAAAATGGGGTGG	1380
Db	1321		
QY	1321	GTGCTGCTACTCCCGAACATCACAGTTTACCAAGTACTGGAAGATCAAAATGGGGTGG	1380
Db	1321		
QY	1381	AATCTAAATCATTAATCTCAGCCTTATGAGAACACCTGCAAAATAAATTAAGAACAAAG	1440
Db	1381		
QY	1381	AATCTAAATCATTAATCTCAGCCTTATGAGAACACCTGCAAAATAAATTAAGAACAAAG	1440
Db	1381		
QY	1441	AAAAATGATATATCTCTTAAGTCTGCTGTAAGAAATGAAGAGTACTTTATGTTTCTCTGAGC	1500
Db	1441		
QY	1441	AAAAATGATATATCTCTTAAGTCTGCTGTAAGAAATGAAGAGTACTTTATGTTTCTCTGAGC	1500
Db	1441		
QY	1501	CAAAGACTCCAGTTAATAAGAACCCAGCATTAAGAGAGAAATCTCACTACGCCCAATCGT	1560
Db	1501		
QY	1501	CAAAGACTCCAGTTAATAAGAACCCAGCATTAAGAGAGAAATCTCACTACGCCCAATCGT	1560
Db	1501		
QY	1561	ACACTACACCTCAAAAGCTAGAAACCAAGTGGCTGCTGAAGAAAACTCCAATTAATAATCCAG	1620
Db	1561		
QY	1561	ACACTACACCTCAAAAGCTAGAAACCAAGTGGCTGCTGAAGAAAACTCCAATTAATAATCCAG	1620
Db	1561		
QY	1621	TAAATTTCAACAGGAAACAGACAAGTTTAATGACAGGTGTCAATTAGCCCTGAGAGCGGTGCC	1680
Db	1621		
QY	1621	TAAATTTCAACAGGAAACAGACAAGTTTAATGACAGGTGTCAATTAGCCCTGAGAGCGGTGCC	1680
Db	1621		
QY	1681	GCTCAGTGGAAATTTGGATCTCAACCAAGCACATATGAGAGAGACTCCCAAAAAGAAAGGAG	1740
Db	1681		
QY	1681	GCTCAGTGGAAATTTGGATCTCAACCAAGCACATATGAGAGAGACTCCCAAAAAGAAAGGAG	1740
Db	1681		
QY	1741	CCAAAGTGTGTTGGAGCCTTGAAAGGGGTGTAAGGTTATCACTGTGCTCACCAGGA	1800
Db	1741		
QY	1741	CCAAAGTGTGTTGGAGCCTTGAAAGGGGTGTAAGGTTATCACTGTGCTCACCAGGA	1800
Db	1741		
QY	1801	GCAAAAGGAAGGTTCTGTCAGAGACGGGCCACAGAGACTTAAAGCTTCACTATAATGTGA	1860
Db	1801		
QY	1801	GCAAAAGGAAGGTTCTGTCAGAGACGGGCCACAGAGACTTAAAGCTTCACTATAATGTGA	1860
Db	1801		
QY	1861	CTACAACTAGATTAGTGAATCCAGATCAAATGTTTGAATGAAATTAATGTTCTTCTTCCAA	1920
Db	1861		
QY	1861	CTACAACTAGATTAGTGAATCCAGATCAAATGTTTGAATGAAATTAATGTTCTTCTTCCAA	1920
Db	1861		
QY	1921	AGAGCATGTTGACTTTGTACAAAGGGTTATACACTGAAGTGTCAAAACACAGTCAAGATT	1980
Db	1921		
QY	1921	AGAGCATGTTGACTTTGTACAAAGGGTTATACACTGAAGTGTCAAAACACAGTCAAGATT	1980
Db	1921		
QY	1981	TTGGGAAAGTGACAATGCAATTTGAATTAAGAAAGTGTGCCAGACTTCAAAAACCCGATGTGG	2040
Db	1981		
QY	1981	TTGGGAAAGTGACAATGCAATTTGAATTAAGAAAGTGTGCCAGACTTCAAAAACCCGATGTGG	2040
Db	1981		
QY	2041	TGGGTATCAGGAGGACGGCTTAAAGGCGATGCTCGGGTTTACAAAAGATTAGTGGAAAG	2100
Db	2041		
QY	2041	TGGGTATCAGGAGGACGGCTTAAAGGCGATGCTCGGGTTTACAAAAGATTAGTGGAAAG	2100
Db	2041		
QY	2101	ACATCCTATCTAGCTGCAAGGTATAATTGATGGATTCTTCCATCTCGCGGATGAGTGTG	2160
Db	2101		
QY	2101	ACATCCTATCTAGCTGCAAGGTATAATTGATGGATTCTTCCATCTCGCGGATGAGTGTG	2160
Db	2101		
QY	2161	GGTGTGATACAGCTACATAAAGACTGTTTATGATCGCTTTTGAATTTTAAAGTTCATTGGAA	2220
Db	2161		
QY	2161	GGTGTGATACAGCTACATAAAGACTGTTTATGATCGCTTTTGAATTTTAAAGTTCATTGGAA	2220
Db	2161		
QY	2221	CTACCAACTTGTGTTTAAAGAGCTATCTTAAAGCAAAATATCTCTTTGTTTAAACAA	2280
Db	2221		
QY	2221	CTACCAACTTGTGTTTAAAGAGCTATCTTAAAGCAAAATATCTCTTTGTTTAAACAA	2280
Db	2221		
QY	2281	GATATTTATTTGTGTATGAATCTTAATCAAGCCCATCTGTCTATGTTTACTGTCTTTTT	2340
Db	2281		
QY	2281	GATATTTATTTGTGTATGAATCTTAATCAAGCCCATCTGTCTATGTTTACTGTCTTTTT	2340
Db	2281		
QY	2341	TAAATCATGTGTTTGTATATTAATTAATTTGTTGACTTTTCTTAGATTCACTTCCATATGTG	2400
Db	2341		
QY	2341	TAAATCATGTGTTTGTATATTAATTAATTTGTTGACTTTTCTTAGATTCACTTCCATATGTG	2400
Db	2341		
QY	2401	AATGTAAGCTCTTAACTATGCTCTCTTTGTTGTAATGTTGTAATTTCTTCTGTAATTAACCAT	2460
Db	2401		

```
|||||
Db 2401 AATGTAAGCTCTTAACATATGCTCTTTGTAATGTAATTTCTTCTGAATAAAACCAT 2460
Qy 2461 TTGTGAATAT 2470
Db 2461 TTGTGAATAT 2470

RESULT 2
US-09-870-937-9
; Sequence 9, Application US/09870937
; Patent No. 6974867
; GENERAL INFORMATION:
; APPLICANT: Wu, Bin
; APPLICANT: Seeley, Todd
; APPLICANT: Williams, Lewis T.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING NEOPLASTIC DISEASE USING
; TITLE OF INVENTION: CHEMOTHERAPY AND RADIATION SENSITIZERS
; FILE REFERENCE: 200130.514/PP-01623.002
; CURRENT APPLICATION NUMBER: US/09/870,937
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 9
; LENGTH: 2470
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-870-937-9

Query Match 100.0%; Score 2470; DB 4; Length 2470;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2470; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TTGGCGGGGGAAGCGGCACCAACCGCGGATCGAAAGATTCTTAGGAACGCGGTACCA 60
Db 1 TTGGCGGGGGAAGCGGCACCAACCGCGGATCGAAAGATTCTTAGGAACGCGGTACCA 60
Qy 61 GCGCGCTCTCTAGGACAGCAGCGCCCTGTCCTTCTGTGCGGCGCGCTCAGCGGTGCC 120
Db 61 GCGCGCTCTCTAGGACAGCAGCGCCCTGTCCTTCTGTGCGGCGCGCTCAGCGGTGCC 120
Qy 121 TCGCCCTCAGGTCTCTTCTTAATCCAAATAAATCTTGAAGAGATGCAAGAT 180
Db 121 TCGCCCTCAGGTCTCTTCTTCTTAATCCAAATAAATCTTGAAGAGATGCAAGAT 180
Qy 181 ATGATGAATCTCTCAATATATGAATATACATGAATCTTGGACAGGTGGCTTTGCAA 240
Db 181 ATGATGAATCTCTCAATATATGAATATACATGAATCTTGGACAGGTGGCTTTGCAA 240
Qy 241 AGGTCAAACTTGCTGCCATATCTTCTTCTGAGAGATGCTATATAAAATCATGGATA 300
Db 241 AGGTCAAACTTGCTGCCATATCTTCTTCTGAGAGATGCTATATAAAATCATGGATA 300
Qy 301 AAAACACACTAGGAGTATTTGCCCGGATCAAAACGAGATGAGGCTTGAAGAACC 360
Db 301 AAAACACACTAGGAGTATTTGCCCGGATCAAAACGAGATGAGGCTTGAAGAACC 360
Qy 361 TGAGACATCAGCATATATGTCAACTCTACATGTGCTAGAGACAGCAACAAATATTC 420
Db 361 TGAGACATCAGCATATATGTCAACTCTACATGTGCTAGAGACAGCAACAAATATTC 420
Qy 421 TGGTTCTTGAGTACTGCTCCGAGGAGCTGTTTGAATATATTAATTTCCAGGATCGCC 480
Db 421 TGGTTCTTGAGTACTGCTCCGAGGAGCTGTTTGAATATATTAATTTCCAGGATCGCC 480
Qy 481 TGTCAAGACAGAGACCGGGTGTCTTCCGTGAGATAGTATCTGCTTGTATGTGC 540
Db 481 TGTCAAGACAGAGACCGGGTGTCTTCCGTGAGATAGTATCTGCTTGTATGTGC 540
Qy 541 ACAGCCAGGGCTATGCTCAGCGGACTCAAGCCAGAAAATTTGCTTTGATGATATC 600
Db 541 ACAGCCAGGGCTATGCTCAGCGGACTCAAGCCAGAAAATTTGCTTTGATGATATC 600
```

```
Qy 601 ATAAATTAAGCTGATGATCTTTGGTCTCTGTGCAAAACCCAGGTAACAAGATTACC 660
Db 601 ATAAATTAAGCTGATGATCTTTGGTCTCTGTGCAAAACCCAGGTAACAAGATTACC 660
Qy 661 ATCTACAGACATGCTGTGGAGTCTGGCTTATGACGACCTGATTAATACAGGCAAT 720
Db 661 ATCTACAGACATGCTGTGGAGTCTGGCTTATGACGACCTGATTAATACAGGCAAT 720
Qy 721 CATATCTTGGATCAGAGGAGATGTTTGGAGCATGGGCATCTGTTATATGTTCTTATGT 780
Db 721 CATATCTTGGATCAGAGGAGATGTTTGGAGCATGGGCATCTGTTATATGTTCTTATGT 780
Qy 781 GTGGATTTCTACCATTTGATGATGATGATGATGATGATGATGATGATGATGATGATG 840
Db 781 GTGGATTTCTACCATTTGATGATGATGATGATGATGATGATGATGATGATGATGATG 840
Qy 841 GAAATATATGATGTTCCCAAGTGGCTCTCCAGTAGCATCTGCTTCTTCAACAATGC 900
Db 841 GAAATATATGATGTTCCCAAGTGGCTCTCCAGTAGCATCTGCTTCTTCAACAATGC 900
Qy 901 TGCAGGTGACCCAAAGAAACGGATTCTATGAAAAATCTATTGAACCATCCCTGGATCA 960
Db 901 TGCAGGTGACCCAAAGAAACGGATTCTATGAAAAATCTATTGAACCATCCCTGGATCA 960
Qy 961 TGCAGGATTAACAATCTATCTGTTGAGTGCAAGCAAGATCCTTTATATCACTCGATG 1020
Db 961 TGCAGGATTAACAATCTATCTGTTGAGTGCAAGCAAGATCCTTTATATCACTCGATG 1020
Qy 1021 ATGATTTGGCTAACAGAACTTTCTGTACATCAGAAACCAAGCAAGCAATGGAGATT 1080
Db 1021 ATGATTTGGCTAACAGAACTTTCTGTACATCAGAAACCAAGCAAGCAATGGAGATT 1080
Qy 1081 TAATTTCACTGTGGCAGTATGATCACTCAGCGCTACCTATCTTCTGCTTCTAGCAAGA 1140
Db 1081 TAATTTCACTGTGGCAGTATGATCACTCAGCGCTACCTATCTTCTGCTTCTAGCAAGA 1140
Qy 1141 AGGCTCGGGGAAAACAGTTCGTTTAAAGCTTCTTCTTCTCTGTTGCAAGCCAGTG 1200
Db 1141 AGGCTCGGGGAAAACAGTTCGTTTAAAGCTTCTTCTTCTCTGTTGCAAGCCAGTG 1200
Qy 1201 CTACCCCAATTCACAGACATCAAGTCAAAATAATGGAAGTCTGGAAGTGAACCAAGTG 1260
Db 1201 CTACCCCAATTCACAGACATCAAGTCAAAATAATGGAAGTCTGGAAGTGAACCAAGTG 1260
Qy 1261 ATAAAAATTTATGTGGCGGATTAATAGACTATGATGATGATGATGATGATGATGATGAT 1320
Db 1261 ATAAAAATTTATGTGGCGGATTAATAGACTATGATGATGATGATGATGATGATGATGAT 1320
Qy 1321 GTGCTGCTACTCTCCCGAATCATCAGTTCACAGTTCACAGTTCACAGTTCACAGTTCAG 1380
Db 1321 GTGCTGCTACTCTCCCGAATCATCAGTTCACAGTTCACAGTTCACAGTTCACAGTTCAG 1380
Qy 1381 AATCTAAATCATTAATCTCAGCTTATGCAAGCAACCTGCAAAATAATTAAGAAACAAG 1440
Db 1381 AATCTAAATCATTAATCTCAGCTTATGCAAGCAACCTGCAAAATAATTAAGAAACAAG 1440
Qy 1441 AAAATGTATATACTCTTAAGTCTGCTGTAAGAAATGAAGAGTACTTATGTTTCTGAGC 1500
Db 1441 AAAATGTATATACTCTTAAGTCTGCTGTAAGAAATGAAGAGTACTTATGTTTCTGAGC 1500
Qy 1501 CAAAGACTCCAGTTTAAAGAACAGCATTAAGAGAGAAATACTCACTACGCGCAATCGTT 1560
Db 1501 CAAAGACTCCAGTTTAAAGAACAGCATTAAGAGAGAAATACTCACTACGCGCAATCGTT 1560
Qy 1561 ACCTACACCTCAAAAGCTAGAACACAGTCCCTGAAAGAACTCCAATTAATAATACAG 1620
Db 1561 ACCTACACCTCAAAAGCTAGAACACAGTCCCTGAAAGAACTCCAATTAATAATACAG 1620
Qy 1621 TAAATTTCAACAGGAAACAGCAAGTTAATGACAGGTGTCATTTAGCCCTGAGAGCGGTGCC 1680
Db 1621 TAAATTTCAACAGGAAACAGCAAGTTAATGACAGGTGTCATTTAGCCCTGAGAGCGGTGCC 1680
Qy 1681 GCTCAGTGGAAATTTGGATCTCAACCAAGCACATATGGAAGAGATCTCAAAAAAGAGGAG 1740
```

```
Db 1681 GCTCAGTGAATTTGGATCTCAACCAAGCACATATGAGGAGACTCCAAAAGAAAGGGAG 1740
Qy 1741 CCNAAGTGTGGAGCCTTGAAGGGGTTGGATAAGGTATCATCTGTGCTCACCCAGA 1800
Db 1741 CCNAAGTGTGGAGCCTTGAAGGGGTTGGATAAGGTATCATCTGTGCTCACCCAGA 1800
Qy 1801 GCAAAAGGAAGGGTCTGCCAGAGACGGCCCAAGAGACTAAAGCTTCACTATAATGTCA 1860
Db 1801 GCAAAAGGAAGGGTCTGCCAGAGACGGCCCAAGAGACTAAAGCTTCACTATAATGTCA 1860
Qy 1861 CTACAACTAGATTAGTGAATCCAGATCAACTGTGTGAATGAATAATGTCTATTTCCAA 1920
Db 1861 CTACAACTAGATTAGTGAATCCAGATCAACTGTGTGAATGAATAATGTCTATTTCCAA 1920
Qy 1921 AGAAGCATGTGACTTTGTACAAAAGGGTTATACACTGAAGTGTCAACACAGTCAATT 1980
Db 1921 AGAAGCATGTGACTTTGTACAAAAGGGTTATACACTGAAGTGTCAACACAGTCAATT 1980
Qy 1981 TTGGGAAAGTGACAAATGCAATTTGAATTAGAAGTGTCCAGCTTCAAAAACCCGATGTG 2040
Db 1981 TTGGGAAAGTGACAAATGCAATTTGAATTAGAAGTGTCCAGCTTCAAAAACCCGATGTG 2040
Qy 2041 TGGGTATCAGGAGCGAGCGCTTAAGGGCGATGCTCTGGGTTTACAAAAGATTAGTGAAG 2100
Db 2041 TGGGTATCAGGAGCGAGCGCTTAAGGGCGATGCTCTGGGTTTACAAAAGATTAGTGAAG 2100
Qy 2101 ACATCTATCTAGCTCAAGGTATTAATGATGGAATCTTCCATCTCCCGGATGATGTG 2160
Db 2101 ACATCTATCTAGCTCAAGGTATTAATGATGGAATCTTCCATCTCCCGGATGATGTG 2160
Qy 2161 GGTGTGATACAGCTACATAAGACACTTATGATGCTTTGATTTTAAAGTTCATTGGAA 2220
Db 2161 GGTGTGATACAGCTACATAAGACACTTATGATGCTTTGATTTTAAAGTTCATTGGAA 2220
Qy 2221 CTACCAACTTGTCTTAAAGAGCTATCTTAAAGCAATATCTCTTTGTTTTTAAACAAA 2280
Db 2221 CTACCAACTTGTCTTAAAGAGCTATCTTAAAGCAATATCTCTTTGTTTTTAAACAAA 2280
Qy 2281 GATATATTTGTGTATGATCTAAATCAAGCCCATCTGCTATATGTTACTGTCTTTT 2340
Db 2281 GATATATTTGTGTATGATCTAAATCAAGCCCATCTGCTATATGTTACTGTCTTTT 2340
Qy 2341 TAATCATGTGTTTGTATATTAATATGTTGACTTTCTAGATCACTTCCATATGTG 2400
Db 2341 TAATCATGTGTTTGTATATTAATATGTTGACTTTCTAGATCACTTCCATATGTG 2400
Qy 2401 AATGTAAGCTTTAACTATGCTCTTTGTAATGTGTAATTTCTTTCTGAAATAAAACCAT 2460
Db 2401 AATGTAAGCTTTAACTATGCTCTTTGTAATGTGTAATTTCTTTCTGAAATAAAACCAT 2460
Qy 2461 TTGTGAATAT 2470
Db 2461 TTGTGAATAT 2470
```

```
RESULT 3
US-09-513-999C-11495
; Sequence 11495, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclert, A.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; FILE REFERENCE: 59.US2.REG
; CURRENT APPLICATION NUMBER: US/09/513,999C
; CURRENT FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/122,487
; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent.pm
```

```
; SEQ ID NO 11495
; LENGTH: 245
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-513-999C-11495

Query Match      8.7%; Score 214; DB 3; Length 245;
Best Local Similarity 95.7%; Pred. No. 1.4e-48;
Matches 220; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Qy 19 CACAACCCGGCGATCGAAAAGATTCTTAGGAACGCGCTACCAAGCCGCTCTCTCAGGACA 78
Db 2 CACAACCCGGCGATCGAAAAGATTCTTAGGAACGCGCTACCAAGCCGCTCTCTCAGGACA 61
Qy 79 GCAGGCCCTGTCCTTCTGTGGGGCGCTCAGCCGCTCCGCGCCCTCAGGTTCTT 138
Db 62 GCAGGCCCTGTCCTTCTGTGGGGCGCTCAGCCGCTCCGCGCCCTCAGGTTCTT 121
Qy 139 TTTCTAATTTCCAAATTAATACTTGCAGAGGACTATGAAGAATTATGATGAATTTCTCAAT 198
Db 122 TTTCTAATTTCCAAATTAATACTTGCAGAGGACTATGAAGAATTATGATGAATTTCTCAAT 181
Qy 199 ATTATGAATTACATCAAACTATTGGACAGGTGGCTTTGCAAGGTCAAA 248
Db 182 ATTATGAATTACATCAAACTATTGGACAGGTGGCTTTGCAAGGTCAAA 231

RESULT 4
US-09-949-016-4303
; Sequence 4303, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4303
; LENGTH: 1863
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-4303

Query Match      8.4%; Score 206.8; DB 3; Length 1863;
Best Local Similarity 55.4%; Pred. No. 3.6e-46;
Matches 469; Conservative 0; Mismatches 362; Indels 15; Gaps 3;
```

```
Qy 212 TGAAACTATTGGACAGGTGGCTTTGCAAGGTCAAACTTGCTGCTCATATCTTACTCG 271
Db 86 TGACACGCTGGGGTGGCGACCTTGGCAAGTGAAGTTGGCAACATGATTTGACTGG 145
Qy 272 AGAGATGGTAGCTATAAAAAATCATGATAAAACACACTAGGGAG-----TGATTTGCC 325
Db 146 GCATAAAGTAGCTGTGAAGATACTCAATCGACAGAAGATTTCGGAGCCTTGATGTGTAGG 205
Qy 326 CCGGATCAAAACGGAGATTGAGCCTTGNAGAACCTGAGACATCAGCATATATGCAACT 385
Db 206 AAAAAATCCGACAGAGAAATTCAGAACTCAAGCTTTTCAGGCATCTCATATAATTAAC 265
Qy 386 CTACCATGTGTAGACAGACAGCCCAAAAAATATTTCATGGTTCTTGAGTACTGCCCTGGAG 445
Db 266 GTACACAGTTCATCAGTACACCATCTGATATTTTCATGGTGTGGAATATGTCTCAGGAG 325
Qy 446 AGAGCTGTTTGTACTATATAATTTCCAGGATCGCTGTCCAGAGAGAGACCGGGTTGT 505
```

Db 326 AGAGCTATTTGATTAATCTGTAAGAAATGAAGCTGGATGAAGAAAGAAAGTGGGGTCT 385  
Qy 506 CTTCCGTCAGATAGTATCTGCTGTGCTTATGTGCAAGCCAGGGCTATGCTCACAGGGA 565  
Db 386 GTTCCAAACAGATCTTCTGCTGTGCTTATGTGCAAGCCAGGGCTATGCTCACAGGGA 445  
Qy 566 CTTCAAGCCAGAAATTTGCTGCTGTGCTTATGTGCAAGCCAGGGCTATGCTCACAGGGA 625  
Db 446 TTTGAAACCTGAAATTTGCTGCTGTGCTTATGTGCAAGCCAGGGCTATGCTCACAGGGA 505  
Qy 626 TCTCTGTGCAAAACCCAGGGTAACAGGATTAACCATCTACAGCATGCTGTGGGAGTCT 685  
Db 506 TCTTTC-----AAACATGATGTCAGATGCTGAAATTTTAAAGCAAGTTGTGGCTCAC 559  
Qy 686 GGCTTATGCAAGCCAGCTGAGTTAATACAGGCAAAATCATATCTTTGGATCAGAGGCAGATGT 745  
Db 560 CAACTATGCTGCACCAAGAGTAATTTTCAGGAAGATTGTATGCAAGGCCAGAGGTAGATAT 619  
Qy 746 TTGGAGCATGGGCATAGTATGTTATGTTCTTATGTGCTGAAATTTTACCAATTTGATGATGA 805  
Db 620 ATGGAGCAGTGGGGTTATTTCTCTATGCTTTTATTTATGTTGGAACCTTTCCATTTTGAATGA 679  
Qy 806 TAATGTAATGGCTTTTATACAGGAAGATTATGAGAGGAATATGATGTTCCCAAGTGGCT 865  
Db 680 CCAATGTCMACTCTTTTAAAGGAATATGATGAGGATCTTCTACCCCTCAATATTT 739  
Qy 866 CTCCTCCAGTAGCATCTGCTTCTTCAACAAATGCTGCAGGTGGAACCCAAAGAAACGGAT 925  
Db 740 AATCCTCTGCTGATGAGCTTTTGAACATATGCTGCAGGTGATCCCAAGAGGGC 799  
Qy 926 TTTATGAAATCTATTTGAACCATCCCTGGATCATGCAAGA---TTAACAATATCCTGT 982  
Db 800 CACAATCAAGATATCAGGGAACATGAATGGTTTAAACAGGACCTTTCCAAATATCTCT 859  
Qy 983 TGAGTGGCAAGCAAGAACTTTTATTCACCTCGATGATGTTGGTGAACAGAACTTTC 1042  
Db 860 TCTGAGGATCCATCATATAGTTCAACCATGATGATGATGAAGCCTTAAAGAAAGTATG 919  
Qy 1043 TGTACA 1048  
Db 920 TGAAGA 925

## RESULT 5

US-09-949-016-2384  
; Sequence 2384, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 2384  
; LENGTH: 2224  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-2384

Query Match 8.0%; Score 197.6; DB 3; Length 2224;  
Best Local Similarity 52.4%; Pred. No. 1.3e-43;  
Matches 487; Conservative 0; Mismatches 434; Indels 9; Gaps 2;

Qy 214 AAACATTTGGGACAGGTGGCTTTGCAAGGTCAAACTTGCCTGCCATATCTCTTACTGGAG 273  
Db 71 AGACCAATTTGGCAAGGGTAATTTTGGCAAGTGAAGTTGGCCGACACATCTCTGACGGGA 130  
Qy 274 AGATGCTAGCTATATAAAATCATGGATATAAAACACACTAGGGAGTGAATTTGCCCGCATCA 333  
Db 131 AAGAGGTAGCTGTGAAGATCATTTGACAGACTCAACTGAATCTCTCCAGCCTCCAGAAAC 190  
Qy 334 AAACG---GAGATTGAGGCTTTGAAGAACCTGAGACATCAGCATATATGCTCAACTCTACC 390  
Db 191 TATTCGGGAACTGAAGATAATGAAGGTTTGAATCATCCCAACATAGTTAAATATTTTG 250  
Qy 391 ATGCTCTAGAGACGCCAACAAATATTCATGGTCTTGGAGTACTGCCCTGGAGGAGC 450  
Db 251 AAGTATTGAGACTGAGAAACCGCTTACCTTGTATGGAGTACGCTAGTGGCGGAGAGG 310  
Qy 451 TGTGTTGACTATATAATTTCCAGGATCGCTCTCAGAGAGGAGAGCCCGGGTTGCTCTCC 510  
Db 311 TATTTGATTACCTAGTGGCTCATGGCAGGATGAAGAAAGAGGCTCGAGCCAAATTC 370  
Qy 511 GTCAGATAGTATCTGCTGTGCTTATGTGCAAGCCAGGGCTATGCTCACAGGACCTCA 570  
Db 371 GCCAGATAGTGTCTGCTGTGAGTACTGTCCACAGAAAGTTTATTTGCCATAGAGACTTAA 430  
Qy 571 AGCCAGAAATTTGCTGCTGTTGATGATATCATATAAATTAAGCTGATGATGCTCTCT 630  
Db 431 AGCCAGAAACCTGCTCTTGGATGCTGATGAATCAAGATTCAGACTTTGGCTTCA 490  
Qy 631 GTGCAAAACCCAGGGTAACAGGATTAACATCTACAGACATGCTGTGGAGTCTGGCTT 690  
Db 491 GCAATGAATC-----ACCTTTGGGAACAAGCTGGACACCTTCTGTGGCAGTCCCCCT 544  
Qy 691 ATSCAGACCTGAGTTAATAACAGGCAATCATATCTTTGGATCAGAGGAGGATTTTGA 750  
Db 545 ATGCTGCCCCAGAACTCTTCCAGGSCAAATAATATGATGACCCCGAGGTGGATGTGGA 604  
Qy 751 GCATGGGCATCTGTTATATGTTCTTATGTGGAATTTCTACCATTTGATGATGATG 810  
Db 605 GCCTAGGATTTATCTCTATACACTGGTCAAGGATCCCTGCTTTTGTGATGACAGAAC 664  
Qy 811 TAATGGCTTTATACAGGAAGATTATGAGAGGAAATATGATGTTCCCAAGTGGCTCTCTC 870  
Db 665 TCAGAGGCTGGGGAACGGTACTGAGGGGAAATACCGTATTCATCTACATGTCCA 724  
Qy 871 CCAGTAGCATTTCTGCTTCTTCAACAAATCTGCTGAGTGGACCCAAAGAAACCGATTTCTA 930  
Db 725 CGGACTGTGAAACCTGCTTAAAGAAATTTCTCATTTAAATCCCAAGAGAGGACACTT 784  
Qy 931 TGAATAATCTATTGAACCATCCCTGGATCATGCAAGATTACAACTATCTCTGTTGAGTGGC 990  
Db 785 TAGACAAATCATGAAGATCGATGGATGAATGTGGGTCAAGAGATGATGAACATAAAGC 844  
Qy 991 AAAGCAAGAACTCTTTTATTCACCTCGATGATGTTGCTTAACAGAACTTTCTGTACATC 1050  
Db 845 CTTAGCTGGAGCACTCCCTGACTCAAGAGCCCGGGGAGACAGAGCTGATGGTGTCCA 904  
Qy 1051 ACAGAAACCAAGGCAAAACAATGGAGGATTTAAATTTCACTGTGGCAGTATGATCACTCA 1110  
Db 905 TGGGTTATACAGGGAAGAGATCCAGGACTCGCTGTGGGCCAGAGATACACAGAGGTGA 964  
Qy 1111 CGGCTACCTATCTTCTGCTTTCTAGCCAGA 1140  
Db 965 TGGCCACCTATCTGCTCTCTGGGCTACAAGA 994

## RESULT 6

US-09-949-016-2385  
; Sequence 2385, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF

FILE REFERENCE: CL001307  
CURRENT APPLICATION NUMBER: US/09/949,016  
CURRENT FILING DATE: 2000-04-14  
PRIOR APPLICATION NUMBER: 60/241,755  
PRIOR FILING DATE: 2000-10-20  
PRIOR APPLICATION NUMBER: 60/237,768  
PRIOR FILING DATE: 2000-10-03  
PRIOR APPLICATION NUMBER: 60/231,498  
PRIOR FILING DATE: 2000-09-08  
NUMBER OF SEQ ID NOS: 207012  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 2385  
LENGTH: 2224  
TYPE: DNA  
ORGANISM: Human  
US-09-949-016-2385

Query Match 8.0%; Score 197.6; DB 3; Length 2224;  
Best Local Similarity 52.4%; Pred. No. 1.3e-43;  
Matches 487; Conservative 0; Mismatches 434; Indels 9; Gaps 2;

QY 214 AAACATTTGGGACAGGTGGCTTTGCAAGGTCAAACTTGCCTGCCATATCTTACTGGAG 273  
DB |||||  
QY 71 AGACCATTTGGCAAGGTAAATTTGGCAAGGTGAAGTTGGCCCGACACATCTCTGACTGGGA 130  
DB |||||  
QY 274 AGATGGTAGCTATAAAAAATCATGCATATAAAACACACTAGGGAGTGATTTGCCCGGATCA 333  
DB |||||  
QY 131 AAGAGGTAGCTGTGAAGATCATTTGACAGACTCACTCACTGAACTCTCCAGCCTCCAGAAC 190  
DB |||||  
QY 334 AAACG---GAGATTGAGGCTTGAAGAACTGAGACATCAGCATATATGTCAACTTACC 390  
DB |||||  
QY 191 TATTCGCGAAGTAAGAAATTAATGAAGTTTGAATCATCCACATAGTAAATATTATTTG 250  
DB |||||  
QY 391 ATGTGTAGAGACAGCCAAATATTCATGTCTTTGAGTACTGCCCTGGAGGAGAGC 450  
DB |||||  
QY 251 AAGTGTATGAGACTGAGAAACCTCTACCTTGTCTCATGTGAGTACGCTAGTGGCGGAGAGG 310  
DB |||||  
QY 451 TGTTTCACTATATAATTTCCAGGATCGCTGTGCAAGAGGAGAGCCCGGTTGTCTTCC 510  
DB |||||  
QY 311 TATTTGATTACCTAGTGGCTCTAGCGCAGGTGAAGAAAGAGGCTCGAGCCAAATTC 370  
DB |||||  
QY 511 GTCAGATAGTATCTGTGTGCTTATGTGACAGCCAGGCTATGCTCAGGAGACTCA 570  
DB |||||  
QY 371 GCCAGATAGTGTCTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 430  
DB |||||  
QY 571 AGCCAGAAATTTGCTGTTTGTGATATCATATAATTAAGCTGATTTGCTGCTCT 630  
DB |||||  
QY 431 AGGCGAAACCTGCTTGGATGCTGATATGAACATCAAGATTGAGACTTTTGGCTTCA 490  
DB |||||  
QY 631 GTCAGAAACCCAGGGTAAACAGGATTACCATCTACAGACATGCTGTGGGAGTCTGGCTT 690  
DB |||||  
QY 491 GCAATGATTC-----ACCTTTGGGAAACAGCTGGACACCTTCTGTGGCAGTCCCCCTT 544  
DB |||||  
QY 691 ATGCAGACCTGATTTAATAAAGGCAAAATCATATCTTGGATCAGAGGAGGATTTTGA 750  
DB |||||  
QY 545 ATGCTGCCCCAGAACTCTTCCAGGCAAAAAATATGATGACCCGAGGTGATGTGGA 604  
DB |||||  
QY 751 GCATGGGCATCTGTTATATGTTCTTATGTGTGATTTTACCATTTGATGATGAATG 810  
DB |||||  
QY 605 GCCTAGGAGTTATCTCTATACACTGGTTCAGCGGATCCCTGCTTTTGTGAGCAGAAC 664  
DB |||||  
QY 811 TAATGGCTTTATACAGAGATTTATGAGAG3AAATATGATGTTCCCAAGTGGCTCTCTC 870  
DB |||||  
QY 665 TCAGAGAGCTGCGGAGACGGTACTGAGGGGAAATACCGTATTTCCATCTACATGTCCA 724  
DB |||||  
QY 871 CCAGTAGCATTTGCTTCTTCAACAAATGTGAGGTGGACCCAAAGAAACGGAATTTCTA 930  
DB |||||  
QY 725 CGGACTGTGAAAACCTGCTTAAAGAAATTTCTCATTTCTTAATCCAGCAGAGAGGCACTT 784  
DB |||||  
QY 931 TGAABAATCTATTGAAACCATCCCTGATCATGCAAGATTACAATCATCTGTTGAGTGGC 990  
DB |||||  
QY 785 TAGAGCAAAATCATGAAAGATCGATGGATGAATGTGGGTACGAGATGATGAACCTAAGC 844  
DB |||||

RESULT 7  
US-09-949-016-1546  
Sequence 1546, Application US/09949016  
Patent No. 6812339  
GENERAL INFORMATION:  
APPLICANT: VENTER, J. Craig et al.  
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
FILE REFERENCE: CL001307  
CURRENT APPLICATION NUMBER: US/09/949,016  
CURRENT FILING DATE: 2000-04-14  
PRIOR APPLICATION NUMBER: 60/241,755  
PRIOR FILING DATE: 2000-10-20  
PRIOR APPLICATION NUMBER: 60/237,768  
PRIOR FILING DATE: 2000-10-03  
PRIOR APPLICATION NUMBER: 60/231,498  
PRIOR FILING DATE: 2000-09-08  
NUMBER OF SEQ ID NOS: 207012  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 1546  
LENGTH: 2950  
TYPE: DNA  
ORGANISM: Human  
US-09-949-016-1546

Query Match 8.0%; Score 197.6; DB 3; Length 2950;  
Best Local Similarity 52.4%; Pred. No. 1.5e-43;  
Matches 487; Conservative 0; Mismatches 434; Indels 9; Gaps 2;

QY 214 AAACATTTGGGACAGGTGGCTTTGCAAGGTCAAACTTGCCTGCCATATCTTACTGGAG 273  
DB |||||  
QY 478 AGACCATTTGGCAAGGTAAATTTGGCAAGGTGAAGTTGGCCCGACACATCTCTGACTGGGA 537  
DB |||||  
QY 274 AGATGGTAGCTATAAAATCATGATTAACACACTAGGAGTGATTTGGCCCGGATCA 333  
DB |||||  
QY 538 AAGAGGTAGCTGTGAAGATCATTTGACAAAGACTCAAACTGAACTCCCTCCAGCCTCCAGAAC 597  
DB |||||  
QY 334 AAACG---GAGATTGAGGCTTTGAAGAACCTGAGACATCAGCATATATGTCAACTCTACC 390  
DB |||||  
QY 598 TATTCGCGAAGTGAATAATGAAGTTTGAATCATCCCAACATAGTAAATATTATTG 657  
DB |||||  
QY 391 ATGTGCTAGAGACAGCCAAACAAATATTCATGGTCTTGGAGTACGCTGAGGAGAGC 450  
DB |||||  
QY 658 AAGTGATTGAGACTGAGAAACGCTCTACTTGTCTATGTAGTACGCTAGTGGCGGAGAGG 717  
DB |||||  
QY 451 TGTTTGACTATATAATTTCCAGGATCGCTGTGAGAGAGGAGGAGCCCGGTTGTCTTCC 510  
DB |||||  
QY 718 TATTTGATTACCTAGTGGCTCATGGCAGGATGAAGAAAGAGGCTCGAGCCAAATCTC 777  
DB |||||  
QY 511 GTCAGATAGTATCTGCTGTGCTTATGTGCACAGCCAGGCTATGTCTCAGAGGAGCTCA 570  
DB |||||  
QY 778 GCAGATAGTGTCTGCTGTGCTGAGTACTGTCCACAGAGTTTATTTGCCATAGAGACTTAA 837  
DB |||||  
QY 571 AGCCAGAAATTTGCTGTTTGTGATGAATATCAATAATTAAGCTGATTTGCTGCTCT 630  
DB |||||  
QY 838 AGGCGAATAACCTGCTCTTGGATGCTGATGAACATCAAGATTGACAGACTTTGGCTTCA 897  
DB |||||  
QY 631 GTGCAAAACCCAGGGTAAACAGGATTACCATCTACAGACATGCTGTGGAGTCTGGCTT 690  
DB |||||  
QY 898 GCAATGAATTC-----ACCTTTGGGAAACAGCTGGACACCTTCTGTGGCAGTCCCCCTT 951  
DB |||||







QY 991 AAAGCAAGAAATCCTTTTATTTCACTCGATGATGATGCGTAACAGAACTTTCTGTACATC 1050  
DB 944 CTTACGTGGAGCACTCCCTGACTACAAGGACCCCGCGGACAGAGCTGATGGTGTCCA 1003  
QY 1051 ACAGAAACACAGCGCAAAACAAATGGAGATTTAATTTTCACTGTGCGAGTATGATCACCTCA 1110  
DB 1004 TGGGTTATACACGGGAAGAGATCCAGACTGCTGGTGGGCCAGAGATACAAGAGGTGA 1063  
QY 1111 CGGCTACCTATCTTCTGCTTCTTAGCCAAGA 1140  
DB 1064 TGGCCACCTATCTGCTTCTGGGCTACAAGA 1093

RESULT 11  
US-10-760-407-1  
; Sequence 1, Application US/10760407  
; Patent No. 6930173  
; GENERAL INFORMATION:  
; APPLICANT: YAN, Chunhua et al.  
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC  
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES  
; FILE REFERENCE: CL001306-DIV II  
; CURRENT APPLICATION NUMBER: US/10/760,407  
; CURRENT FILING DATE: 2004-01-21  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1  
; LENGTH: 2175  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-760-407-1

Query Match 7.9%; Score 196; DB 3; Length 2175;  
Best Local Similarity 52.3%; Pred. No. 3 6e-43;  
Matches 486; Conservative 0; Mismatches 435; Indels 9; Gaps 2;

QY 214 AAACATATGGGACAGTGGCTTTGCAAGGTCAAACTTGCCTGCCATATCTTACTGGAG 273  
DB 170 AGACCATTTGGCAAGGTAAATTTTGGCAAGTGAAGTTGGCCGACACATCTGACTGGGA 229  
QY 274 AGATGGTAGCTATAAAATCATGGATGAATAAAACACACTAGGGAGTATTTGCCCGGATCA 333  
DB 230 AAGAGGTAGCTGTGAAGATCATTTGCAAGACTCAACTGAACCTCTCCAGCTCCAGAAAC 289  
QY 334 AAACG---CAGATTGAGGCTTTGAAGAACTTGAGACATCAGCATATATGCTCACTTACC 390  
DB 290 TATTCGGGAGTAGAATAAATGAGGTTTGAATCATCCCAACATAGTTAAATTTATTTG 349  
QY 391 ATGTGCTAGACAGCAACAAATATTTTCAATGTTCTTGAAGTACTGCCCTGGAGGAGAGC 450  
DB 350 AAGTATTGAGACTGAGAAACGCTTACTTGTCTATGAGTACGCTAGTGGCGGAGG 409  
QY 451 TGTTTGACTATATAATTTCCAGGATCGCTGTGCAAGAGGAGACCCGGGTTGCTTCC 510  
DB 410 TATTTGATTACCTAGTGGCTCATGGCAGGATGAAGAAAGAGGCTCGAGCCAAATTC 469  
QY 511 GTCAGATAGTATCTGCTGTTGCTTATGTGTCAGAGCCAGGCTATGCTCACAGGAGCTCA 570  
DB 470 GCCAGTAGTGTCTGCTGTGCGAGTACTGTCAACAGAAAGTTATTTGTCATAGAGACTTAA 529  
QY 571 AGCCAGAAATTTGCTGTTTGTATGAATATATCAATAATTAAGCTGATTTGACTTTGCTCT 630  
DB 530 AGGCAGAAACCTGCTTGGATGCTGATGATGAATCAAGATTCAGAGCTTGGCTTCA 589  
QY 631 GTGCAAAACCCAGGGTAAACAGGATTAACATCTACAGACATGCTGTGGAGATCTGGCTT 690  
DB 590 GCAATGAATTC-----ACCTTTGGGAAACAAGCTGGACACCTTCTGTGGCAGTCCCCCTT 643  
QY 691 ATGCAGCACTGATTAATACAGCAATATATCTTGGATCAGAGGACAGATTTTGA 750  
DB 644 ATGCTGCCCCAGAACTCTTCCAGGGGCAAAATATATGATGGACCCGAGGTGGATGTGGA 703

QY 751 GCAATGGCATACTGTTTATATGTTCTTATGTTGGATTTTCTACCAATTTGATGATAATG 810  
DB 704 GCCTAGAGCTTATCTCTATACACTGGTCAAGGATCCCTGCTTTTGTGGACAGAAC 763  
QY 811 TAATGCTTTTATACAAGAGATTTAGAGAGGAATATGATGTTCCCAAGTGGCTCTCTC 870  
DB 764 TCAAGAGAGCTGGGGAGCGGCTACTGAGGGGAAATACCGTATTTCCATTTCTACATGTCCA 823  
QY 871 CCAGTAGCATTTCTGCTTCTTCAACAAATGCTGAGGTGGACCCCAAGAAACCGATTTCTA 930  
DB 824 CGGACTGTGAARACCTGCTTTAAGAAATTTCTCATTTTAAATCCACAGAGAGGCACTT 883  
QY 931 TGAATAATCTATTGAACCATCCCTGATCATGCAAGATTAACAATCTCTGTTGAGTGGC 990  
DB 884 TAGAGCAATCATGAAGATCGATGGTGAATGTGGTCAAGAGATGATGAACATAAGC 943  
QY 991 AAGCAAGAAATCTTTTATTTCACTCGATGATGATTCGTAACAGAACTTTCTGTACATC 1050  
DB 944 CTTACGTGGAGCACTTCCCTGACTACAAGGACCCCGCGGACAGAGCTGATGGTGTCCA 1003  
QY 1051 ACAGAAACACAGGCAACAAACAAATGGAGGATTTAAATTTCACTGTGGCAGTATGATCACCTCA 1110  
DB 1004 TGGGTTATACACGGGAAGAGATCCAGGACTCGTGGTGGGCCAGAGATACAAGAGGTGA 1063  
QY 1111 CGGCTACCTATCTTCTGCTTCTTAGCCAAGA 1140  
DB 1064 TGGCCACCTATCTGCTTCTGGGCTACAAGA 1093

RESULT 12  
US-09-101-146-44  
; Sequence 44, Application US/09101146  
; Patent No. 6124125  
; GENERAL INFORMATION:  
; APPLICANT: Dartmouth College, St. Vincents Institute of  
; APPLICANT: Medical Research, Kemp et al.  
; TITLE OF INVENTION: No. 6124125el AMP Activated Protein Kinase  
; NUMBER OF SEQUENCES: 64  
; CORRESPONDENCE ADDRESS:  
; ADDRESSER: Jane Massey Licata, Esq.  
; STREET: 66 E. Main Street  
; CITY: Marlton  
; STATE: NJ  
; COUNTRY: USA  
; ZIP: 08053  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.44 Mb STORAGE  
; COMPUTER: IBM PC  
; OPERATING SYSTEM: WINDOWS 95  
; SOFTWARE: WORDPERFECT 6.0 FOR WINDOWS  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/101,146  
; FILING DATE: October 7, 1998  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PN7450  
; FILING DATE: 8 JAN 1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Jane Massey Licata  
; REGISTRATION NUMBER: 32,257  
; REFERENCE/DOCKET NUMBER: DC-0050  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (856) 810-1515  
; TELEFAX: (856) 810-1454  
; INFORMATION FOR SEQ ID NO: 44:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1647  
; TYPE: Nucleic acid  
; STRANDEDNESS: Single  
; TOPOLOGY: Linear  
; ANTI-SENSE: No  
US-09-101-146-44



Query Match	7.3%	Score 181.2;	DB 3;	Length 2210;
Best Local Similarity	56.4%;	Pred. No. 4.2e-39;		
Matches 361;	Conservative 0;	Mismatches 273;	Indels 6;	Gaps 1;
Qy	330	ATCAAAACGAGATTGAGGCCTTGAAGAACTCGAGACATCAGCATATATGTCACCTCTAC	389	
Db	89	ATAAAACGAAATTTCAAATCTAAAACTCTTTGTCATCTCTCATATTTATCAAACTATAC	148	
Qy	390	CATGTGCTAGACAGCGCAACAAATATTTCATGGTTCTTGAGTACTGCCCTGGAGGAGAG	449	
Db	149	CAGGTATCAGCACTCCACAGATTTTTTATGGTAATGGAATATGTGCTGGAGGTGAA	208	
Qy	450	CTGTTTGACTATATAATTTCCCAAGATCGCTGTCAAGAGAGAGACCCGGGTTGCTTC	509	
Db	209	TTATTTGACTACATCTGTAAGCATGACGGGTTGAAGAGATGGAAGCCAGCGCGCTCTTT	268	
Qy	510	CGTCAGATAGTATCTGCTGTTGCTTATATGTCACAGCCAGGCGTATGCTCACAGGACCTC	569	
Db	269	CAGCAGATTTCTGCTGTGGTATTAAGTTCATAGGCATATGGTTGTTCTACAGACCTG	328	
Qy	570	AAGCCAGAAATTTGCTTTTGATGAATATCAATAATTAAGCTGATTGACTTTGGTCTC	629	
Db	329	AAACCCAGAGATGTCTGTTGGATGACACATGAATGCCAAGATAGCCGATTTCCGGATTA	388	
Qy	630	TGTGCAAAACCCAAGGTTAAACAGGATTAACATCTACAGACATGCTGTGGAGTCTGGCT	689	
Db	389	TCT-----ATATGATGTCAGATGGTGAATTTCTGAGAACTAGTTGCGGATCTCCAAT	442	
Qy	690	TATGCAACCTCAGTTTAATCAAGGCAATCATATCTTGGATCAGAGGAGATGTTTGG	749	
Db	443	TATGCAACCTGAAAGTCATCTCAGGCAGATTTGATGCAGGTCCTGAAGTTGATATCTGG	502	
Qy	750	AGCATGGGCATCTGTTATATGTTCTTATGTTGGATTTCTACATTTGATGATGATAT	809	
Db	503	AGCTGTGGTGTATCTTGTATGCTCTCTTTGTGGCAACCCCTCCCATTTGATGATGAGCAT	562	
Qy	810	GTAATGCTTTTATCAAGAGATTTATGAGAGAAATATGATGCTTTCCCAAGTGGCTCTCT	869	
Db	563	GTACCTACGTTATTTAAGAAAGATCCGAGGGGGTGTCTTTTATATCCAGAAATATCTCAAT	622	

Qy	870	CCAGTAGCATTCTGCTTCTTCAACAAATGCTGAGGTGACCCAAAGAAACGGATTCT	929
Db	623	CGTTCTGTGCCCACTCTCTCTGATGCATATGCTGAGTTGACCCACTGAAACGAGCAACT	682
Qy	930	ATGAAAAATCTATTGAACCATCCCTGGATCATGCAAGATT	969
Db	683	ATCAAGACATAAGAGAGCATGAATGGTTTAAACAAGATT	722

Search completed: November 21, 2006, 19:46:41  
Job time : 463 secs

GenCore version 5.1.9  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 21, 2006, 19:34:26 ; Search time 11521 Seconds  
(without alignments)  
11988.608 Million cell updates/sec

Title: US-10-656-598-1  
Perfect score: 2470  
Sequence: 1 tggcgggcggaagcgccca.....ataaacacatttgtgaatat 2470

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 48236798 seqs, 27959665780 residues

Total number of hits satisfying chosen parameters: 96473596

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:\*

1: gb\_est1:\*

2: gb\_est3:\*

3: gb\_est4:\*

4: gb\_est5:\*

5: gb\_est6:\*

6: gb\_estc:\*

7: gb\_est2:\*

8: gb\_est7:\*

9: gb\_est8:\*

10: gb\_est9:\*

11: gb\_gss1:\*

12: gb\_gss2:\*

13: gb\_gss3:\*

14: gb\_gss4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	1615	65.4	1615	6	CR605252 full-leng
2	1548.8	62.7	2469	6	AK145021 Mus muscu
3	1543.4	62.5	2430	6	AK145316 Mus muscu
4	1543.4	62.5	2914	6	AK011932 Mus muscu
5	1541.8	62.4	2447	6	AK167457 Mus muscu
6	1541.8	62.4	2469	6	AK164138 Mus muscu
7	1535.8	62.2	2541	6	AK161029 Mus muscu
8	1525.8	61.8	2396	6	BC020416 Mus muscu
9	1521.4	61.6	2559	6	AK137000 Mus muscu
10	1520.4	61.6	2488	6	AK010498 Mus muscu
11	1368	55.4	1956	14	DQ048096 Homo sapi
12	1036.4	42.0	1096	1	AL577157 AL577157
13	1020.6	41.3	1030	1	AL554470 AL554470
14	872.4	35.3	1091	3	BM928784 AGENCOURT
15	859	34.8	1898	14	DQ048097 Pan trogl
16	834.8	33.8	855	2	BM471174 AGENCOURT
17	830.6	33.6	901	1	AU130934 AU130934
18	822.6	33.3	960	2	BG682772
19	813.2	32.9	852	1	AU124892 AU124892

20	792.6	32.1	1007	3	BQ646763
21	781	31.6	854	1	AU124737 AGENCOURT
22	778	31.5	1006	3	BM927868 AGENCOURT
23	770.6	31.2	945	3	BQ650221 AGENCOURT
24	756	30.6	762	9	CX872457 HESCA_69
25	744.6	30.1	803	2	BI086933 602850361
26	743.2	30.1	754	1	AU130717 AU130717
c	735.8	29.8	839	1	AI693521 w43e08.x
28	735.2	29.8	848	3	BQ221427 AGENCOURT
29	735.2	29.8	1034	2	BG396052 602458792
30	733	29.7	734	4	BX105144 BX105144
31	726.4	29.4	750	3	BU595645 AGENCOURT
32	720.4	29.2	806	2	BG105946 602290134
c	715.8	29.0	882	1	AA767453 n282D06.s
34	714.8	28.9	746	4	CA426797 UI-H-FE1-
c	710.6	28.8	760	4	CA749050 UI-H-FE1-
36	700	28.3	818	9	DR155628 HESC2_56
37	699.8	28.3	862	7	BE881704 601489932
c	690.4	28.0	772	2	BG683434 602651263
39	682.8	27.6	714	4	CA310724 UI-CF-FN0
40	682.4	27.6	818	1	AM036902 AM036902
41	681.4	27.6	733	2	BG286546 602382548
42	679.2	27.5	853	8	CV800888 AGENCOURT
c	678.8	27.5	715	3	BU622206 UI-H-PH1-
44	660	26.7	672	8	CN307609 328781381
45	654.8	26.5	890	2	BG752315 602731280

## ALIGNMENTS

RESULT 1	CR605252	1615 bp	mRNA	linear	HTC 21-JUL-2004
LOCUS	CR605252	full-length cDNA clone CS0DI084YI07 of Placenta Cot 25-normalized			
DEFINITION	CR605252	of Homo sapiens (human).			
ACCESSION	CR605252.1	GI:50486059			
VERSION	CR605252.1	GI:50486059			
KEYWORDS	HTC; CDSLT cDNA.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	1 (bases 1 to 1615)				
AUTHORS	Li, W.B., Gruber, C., Jessee, J. and Polayes, D.				
TITLE	Full-length cDNA libraries and normalization				
JOURNAL	Unpublished				
REMARK	Contact : Feng Liang Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com/ Invitrogen Corporation 1600 Faraday Avenue				
REFERENCE	2 (bases 1 to 1615)				
AUTHORS	Genoscope.				
TITLE	Direct Submission				
JOURNAL	Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr)				
COMMENT	- Web : www.genoscope.cns.fr) 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies, a division of Invitrogen.				
FEATURES	Location/Qualifiers				
source	1..1615				
	/organism="Homo sapiens"				
	/mol_type="mRNA"				
	/db_xref="taxon:9606"				
	/clone="CS0DI084YI07"				
	/tissue type="Placenta Cot 25-normalized"				
	/plasmid="pCMVSPORT_6"				
ORIGIN					
Query Match	65.4%	Score	1615;	DB	6; Length 1615;



Best Local Similarity 100.0%; Pred. No. 0;			
Matches 1615; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
QY	61 GC	CGCGTCTCTCAGGACAGCAGGCCCTGTCCTTCTGTGCGGCGCGCTCAGCGTGGCC	120
Db	1 GC	CGCGTCTCTCAGGACAGCAGGCCCTGTCCTTCTGTGCGGCGCGCTCAGCGTGGCC	60
QY	121 TC	CGCCCTCAGGTTCTTTTCTAAATTC	180
Db	61 TC	CGCCCTCAGGTTCTTTTCTAAATTC	120
QY	181 AT	GATGAACCTCTCAATATATGAATATCAATGAATATGGAACAGGTGGCTTGC	240
Db	121 AT	GATGAACCTCTCAATATATGAATATCAATGAATATGGAACAGGTGGCTTGC	180
QY	241 AG	GTCAAACTTGCTGCTCATCTCTTCTGAGAGATGGTAGCTATAAAATCATGGATA	300
Db	181 AG	GTCAAACTTGCTGCTCATCTCTTCTGAGAGATGGTAGCTATAAAATCATGGATA	240
QY	301 AA	ACACACTAGGAGTGTGTCCTGCGGATCAAAACGGAGATTGAGGCTTGAAGAAC	360
Db	241 AA	ACACACTAGGAGTGTGTCCTGCGGATCAAAACGGAGATTGAGGCTTGAAGAAC	300
QY	361 TG	AGACATCAGCATATATGTCATCACTCTACCATGCTCTAGAGACAGCCAAATATTTCA	420
Db	301 TG	AGACATCAGCATATATGTCATCACTCTACCATGCTCTAGAGACAGCCAAATATTTCA	360
QY	421 TG	CTTCTGAGTACTGCCCTGGAGGAGAGCTGTTTGACTATATATTTCCAGGATCGC	480
Db	361 TG	CTTCTGAGTACTGCCCTGGAGGAGAGCTGTTTGACTATATATTTCCAGGATCGC	420
QY	481 TG	CAGAGAGAGACCGCGGTTGTCTTCGCTCAGATAGTATCTGCTGTTCTTATGTGC	540
Db	421 TG	CAGAGAGAGACCGCGGTTGTCTTCGCTCAGATAGTATCTGCTGTTCTTATGTGC	480
QY	541 AC	AGCAGGCTATGCTCAAGGACCTCAAGCAGAGAAATTTGCTGTTGATGAATATC	600
Db	481 AC	AGCAGGCTATGCTCAAGGACCTCAAGCAGAGAAATTTGCTGTTGATGAATATC	540
QY	601 AT	AAATTAAGCTGATGACTTGGTCTCTGTCGCAAAACCAAGGTAACAGGATTACC	660
Db	541 AT	AAATTAAGCTGATGACTTGGTCTCTGTCGCAAAACCAAGGTAACAGGATTACC	600
QY	661 AT	CTACAGACATGCTGTGGAGTCTGGCTTATGAGCAGCTGAGTTAATCAAGGCAAT	720
Db	601 AT	CTACAGACATGCTGTGGAGTCTGGCTTATGAGCAGCTGAGTTAATCAAGGCAAT	660
QY	721 CA	TATCTTGGATCAGAGGCAGATGTTTGGAGCATGGGCATCTGTTATATGTTCTTATGT	780
Db	661 CA	TATCTTGGATCAGAGGCAGATGTTTGGAGCATGGGCATCTGTTATATGTTCTTATGT	720
QY	781 GT	GGATTTCTACCAATTTGATGATTAATGATGCTTATACAGAGATTAATGAG	840
Db	721 GT	GGATTTCTACCAATTTGATGATTAATGATGCTTATACAGAGATTAATGAG	780
QY	841 GA	AAATATGATGTTCCCAAGTGGCTCTCCAGTAGCATCTGCTTCTTCAACAAATGC	900
Db	781 GA	AAATATGATGTTCCCAAGTGGCTCTCTCCAGTAGCATCTGCTTCTTCAACAAATGC	840
QY	901 TG	CAGTGGACCAAGAAACGATTTCTATGAAAAATCTATTGAAACCATCTCGATCA	960
Db	841 TG	CAGTGGACCAAGAAACGATTTCTATGAAAAATCTATTGAAACCATCTCGATCA	900
QY	961 TG	CAAGTTACACTATCTGTTGAGTGGCAAGCAGATCTTTTATTCACCTCGATG	1020
Db	901 TG	CAAGTTACACTATCTGTTGAGTGGCAAGCAGATCTTTTATTCACCTCGATG	960
QY	1021 AT	GATGCGTAACAGAACTTTCTGTACATCACAGAAACCAAGGCAACCAATGGAGATT	1080
Db	961 AT	GATGCGTAACAGAACTTTCTGTACATCACAGAAACCAAGGCAACCAATGGAGATT	1020
QY	1081 TA	ATTTCACTGTGGCAGTATGATCACTCAAGGCTACCTATCTTCTGCTTCTAGCAAGA	1140

---

Db	1021	TA	ATTTCACTGTGGCAGTATGATCACTCAGCGGTACTCTATCTTCTCTAGCCAAGA	1080
QY	1141	AG	CTCGGGGAAACCCAGTTCGTTTAAAGGCTTTCTTTCTTCTCTGTCGACCAAGCCAGTG	1200
Db	1081	AG	CTCGGGGAAACCCAGTTCGTTTAAAGGCTTTCTTTCTTCTCTGTCGACCAAGCCAGTG	1140
QY	1201	CT	ACCCCATTCACAGACATCAAGTCAAAATTAATTTGAGTCTGGAAGATGTGACCGCAAGTG	1260
Db	1141	CT	ACCCCATTCACAGACATCAAGTCAAAATTAATTTGAGTCTGGAAGATGTGACCGCAAGTG	1200
QY	1261	AT	AAAAAATTTATGTGGCGGATTAATAGACTATGATTTGGTGTGAAGATGATTTATCAACAG	1320
Db	1201	AT	AAAAAATTTATGTGGCGGATTAATAGACTATGATTTGGTGTGAAGATGATTTATCAACAG	1260
QY	1321	GT	GCTGCTACTCCCCGAAACATCACAGTTTACCAAGTACTTGGACAGAAATCAAAATGGGGTGG	1380
Db	1261	GT	GCTGCTACTCCCCGAAACATCACAGTTTACCAAGTACTTGGACAGAAATCAAAATGGGGTGG	1320
QY	1381	AA	TCTAAATCATTAATCTCAGCCTTATGACAGAACCTCTGCAAAATAAATTAAGAACAAAG	1440
Db	1321	AA	TCTAAATCATTAATCTCAGCCTTATGACAGAACCTCTGCAAAATAAATTAAGAACAAAG	1380
QY	1441	AA	ATGTATATCTCTTAAGTCTGCTGTAAAGAAATGAAGTACTTTATGTTTCTCTGAGC	1500
Db	1381	AA	ATGTATATCTCTTAAGTCTGCTGTAAAGAAATGAAGTACTTTATGTTTCTCTGAGC	1440
QY	1501	CA	AGACTCCAGTTAATAGAACACAGCATAAGAGAGAAATCTCACTACGCCAAATCGTT	1560
Db	1441	CA	AGACTCCAGTTAATAGAACACAGCATAAGAGAGAAATCTCACTACGCCAAATCGTT	1500
QY	1561	AC	ACTACACCTTCAAAAGCTAGAAACCACTGCTGAAAGAACTCCAAATTAATAATACCAG	1620
Db	1501	AC	ACTACACCTTCAAAAGCTAGAAACCACTGCTGAAAGAACTCCAAATTAATAATACCAG	1560
QY	1621	TA	AAATTTCAACAGACACAGCAAGTTAATGACAGGTGTCATTTAGCCCTGAGAGCGG	1675
Db	1561	TA	AAATTTCAACAGACACAGCAAGTTAATGACAGGTGTCATTTAGCCCTGAGAGCGG	1615

---

RESULT 2		2469 bp	mRNA	linear	HTC 21-SEP-2005
AK145021		Mus musculus mammary gland RCB-0526 Jyg-MC (A) cDNA, RIKEN			
LOCUS		full-length enriched library, clone:G830026003 product:maternal			
DEFINITION		embryonic leucine zipper kinase, full insert sequence.			
ACCESSION		AK145021			
VERSION		AK145021.1 GI:74202957			
KEYWORDS		HTC; CAP trapper.			
SOURCE		Mus musculus (house mouse)			
ORGANISM		Mus musculus			
		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
		Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;			
		Sciurognathi; Muroidea; Muridae; Murinae; Mus.			
REFERENCE		1			
AUTHORS		Carninci, P. and Hayashizaki, Y.			
TITLE		High-efficiency full-length cDNA cloning			
JOURNAL		Meth. Enzymol. 303, 19-44 (1999)			
PUBMED		10349636			
REFERENCE		2			
AUTHORS		Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,			
TITLE		Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.			
JOURNAL		Normalization and subtraction of cap-trapper-selected cDNAs to			
PUBMED		prepare full-length cDNA libraries for rapid discovery of new genes			
JOURNAL		Genome Res. 10 (10), 1617-1630 (2000)			
PUBMED		11042159			
REFERENCE		3			
AUTHORS		Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,			
		Konno, H., Akiyama, J., Nishi, K., Kitsunai, T., Tashiro, H., Itoh, M.,			
		Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,			
		Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,			
		Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M.,			
		Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J.,			
		Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.			

TITLE JOURNAL PUBMED REFERENCE AUTHORS	RIKEN integrated sequence analysis (RISA) system-394-format sequencing pipeline with 384 multicapillary sequencer Genome Res. 10 (11), 1757-1771 (2000) 11076861	
	4	Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y., Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S., Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamanaka, I., Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R., Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T., Fleishmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H., Kuehl, P., Lewis, S., Matsuo, Y., Nikaide, I., Pesole, G., Kachiwa, H., Quackenbush, J., Schriml, L.M., Staubli, F., Suzuki, R., Tomita, M., Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H., Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N., Carninci, P., De Bonaldo, M.F., Brownstein, M.J., Bult, C., Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D., Hofmann, M., Hume, D.A., Kamiya, M., Lee, N.H., Lyons, P., Marchionni, L., Mashima, J., Mazzarelli, J., Mombaerts, P., Nordone, P., Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H., Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K.P., Suzuki, H., Toyooka, K., Wang, K.H., Weitz, C., Whittaker, C., Wilming, L., Wyshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawai, H., Kohetsuki, S. and Hayashizaki, Y.
	5	RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium Functional annotation of a full-length mouse cDNA collection Nature 409 (6821), 685-690 (2001) 11217851
	6	CONSTRM JOURNAL PUBMED REFERENCE AUTHORS
TITLE JOURNAL PUBMED REFERENCE AUTHORS	Okaaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S., Nikaide, I., Otao, N., Saito, R., Suzuki, H., Yamanaka, I., Kiyosawa, H., Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A., Schonbach, C., Gojobori, T., Baldarelli, R., Hill, D.P., Bult, C., Hume, D.A., Quackenbush, J., Schriml, L.M., Kanapin, A., Matsuda, H., Batalov, S., Beisel, K.W., Blake, J.A., Brad, D., Brusica, V., Chothia, C., Corbani, L.E., Cousins, S., Dalla, E., Dragani, T.A., Fletcher, C.F., Forrest, A., Frazer, K.S., Gaasterland, T., Gariboldi, M., Gissi, C., Godzik, A., Gough, J., Grimmond, S., Gustincich, S., Hirokawa, N., Jackson, I.J., Jarvis, E.D., Kanai, A., Kawai, J., Kawai, Y., Kedesz, R.M., King, B.L., Konagaya, A., Kurochkin, I.V., Lee, Y., Lenhard, B., Lyons, P.A., Maglott, D.R., Maltais, L., Marchionni, L., McKenzie, L., Miki, H., Nagashima, T., Numata, K., Okido, T., Pavan, W.J., Perte, G., Pesole, G., Petrovsky, N., Pillai, R., Pontius, J.U., Qi, D., Ramachandran, S., Ravasi, T., Reed, J.C., Reed, D.J., Reid, J., Ring, B.Z., Ringwald, M., Sandelin, A., Schneider, C., Semple, C.A., Setou, M., Shimada, K., Sultana, R., Takenaka, Y., Taylot, M.S., Teasdale, R.D., Tomita, M., Verardo, R., Wagner, L., Wahlstedt, C., Wang, Y., Watanabe, Y., Wells, C., Wilming, L.G., Wyshaw-Boris, A., Yanagisawa, M., Yang, I., Yang, L., Yuan, Z., Zavolan, M., Zhu, Y., Zimmer, A., Carninci, P., Hayatsu, N., Hirozane-Kitahara, T., Konno, H., Nakamura, M., Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, J., Aizawa, K., Arakawa, T., Fukuda, S., Hara, A., Hashizume, W., Imotani, K., Ishii, Y., Itoh, M., Kigawa, I., Miyazaki, A., Sakai, K., Sasaki, D., Shibata, K., Shinagawa, A., Yasunishi, A., Yoshino, M., Waterston, R., Lander, E.S., Rogers, J., Birney, E. and Hayashizaki, Y.	
	6	PANTOM Consortium Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs Nature 420 (6915), 563-573 (2002) 12466851
	6	CONSTRM JOURNAL PUBMED REFERENCE AUTHORS
	Carninci, P., Kasukawa, T., Katayama, S., Gough, J., Frith, M.C., Maeda, N., Oyama, R., Ravasi, T., Lenhard, B., Wells, C., Kodzius, R., Shinkawa, K., Bajic, V.B., Brenner, S.E., Batalov, S., Forrest, A.R., Zavolan, M., Davis, M.J., Wilming, L.G., Aidinis, V., Allen, J.E., Ambesi-Impombato, A., Apweiler, R., Aturaliya, R.N., Bailey, T.L., Banal, M., Baxter, L., Beisel, K.W., Bersano, T., Bono, H., Chalk, A.M., Chiu, K.P., Choudhary, V., Christoffels, A., Clutterbuck, D.R., Crowe, M.L., Dalla, E., Dalrymple, B.P., de Bono, B., Della Gatta, G., di Bernardo, D., Down, T., Engstrom, P., Fagioli, M., Faulkner, G., Fletcher, C.F., Fukushima, T., Furuno, M., Futaki, S., Gariboldi, M.,	

TITLE JOURNAL PUBMED REFERENCE AUTHORS	Georgii-Hemming, P., Giegeras, T.R., Gojobori, T., Green, R.E., Gustincich, S., Harbers, M., Hayashi, Y., Hensch, T.K., Hirokawa, N., Hill, D., Huminecki, L., Iacono, M., Ikey, K., Iwano, A., Ishikawa, T., Jakt, M., Kanapin, A., Kato, M., Kawasawa, Y., Kelson, J., Kitamura, H., Kitano, H., Kollias, G., Kriehnan, S.P., Kruger, A., Kummerfeld, S.K., Kurochkin, I.V., Lareau, L.F., Lazarevic, D., Lipovich, L., Liu, J., Liuni, S., McWilliam, S., Madan Babu, M., Madera, M., Marchionni, L., Matsuda, H., Matsuzawa, S., Miki, H., Mignone, F., Miyake, S., Morris, K., Mottagui-Tabar, S., Mulder, N., Nakano, N., Nakachi, H., Ng, P., Nilsson, R., Nishiguchi, S., Nishikawa, S., Nori, F., Ohara, O., Okazaki, Y., Orlando, V., Pang, K.C., Pavan, W.J., Pavese, G., Pesole, G., Petrovsky, N., Piazza, S., Reed, J., Reid, J.F., Ring, B.Z., Ringwald, M., Rost, B., Ruan, Y., Salzberg, S.L., Sandelin, A., Schneider, C., Schonbach, C., Sekiguchi, K., Semple, C.A., Seno, S., Sessa, L., Sheng, Y., Shibata, Y., Shimada, H., Shimada, K., Silva, D., Sinclair, B., Sperling, S., Stupka, E., Sugita, K., Sultana, R., Takenaka, Y., Taki, K., Tammoja, K., Tan, S.L., Tang, S., Taylor, M.S., Tegner, J., Teichmann, S.A., Ueda, H.R., van Nimwegen, E., Verardo, R., Wei, C.L., Yagi, K., Yamanishi, H., Zabarovsky, E., Zhu, S., Zimmer, A., Hide, W., Bult, C., Grimmond, S.M., Teasdale, R.D., Liu, E.T., Brusica, V., Quackenbush, J., Wahlstedt, C., Mattick, J.S., Hume, D.A., Kai, C., Sasaki, D., Tomaru, Y., Fukuda, S., Kanamori-Katayama, M., Suzuki, M., Aoki, J., Arakawa, T., Iida, J., Imamura, K., Itoh, M., Kato, T., Kawaji, H., Kawagashira, N., Kawashima, T., Kojima, M., Kondo, S., Konno, H., Nakano, K., Ninomiya, N., Nishio, T., Okada, M., Plessey, C., Shibata, K., Shiraki, T., Suzuki, S., Tagami, M., Waki, K., Watahiki, A., Okamura-Oho, Y., Suzuki, H., Kawai, J. and Hayashizaki, Y.	
	7	PANTOM Consortium The transcriptomic landscape of the mammalian genome Science 309 (5740), 1559-1563 (2005) 16141072
	7	CONSTRM JOURNAL PUBMED REFERENCE AUTHORS
	Katayama, S., Tomaru, Y., Kasukawa, T., Waki, K., Nakanishi, M., Nakamura, M., Nishida, H., Yagi, K., Suzuki, M., Kawai, J., Suzuki, H., Carninci, P., Hayashizaki, Y., Wells, C., Frith, M., Ravasi, T., Pang, K.C., Hallinan, J., Mattick, J., Hume, D.A., Lipovich, L., Batalov, S., Engstrom, P.G., Mizuno, Y., Faghihi, M.A., Sandelin, A., Chalk, A.M., Mottagui-Tabar, S., Liang, Z., Lenhard, B. and Wahlstedt, C.	
TITLE JOURNAL PUBMED REFERENCE AUTHORS	RIKEN Genome Exploration Research Group Antisense transcription in the mammalian transcriptome Science 309 (5740), 1564-1566 (2005) 16141073	
	8	(bases 1 to 2469) Arakawa, T., Carninci, P., Fukuda, S., Hashizume, W., Hayashida, K., Hori, F., Iida, J., Imamura, K., Imotani, K., Itoh, M., Kanagawa, S., Kawai, J., Kojima, M., Konno, H., Murata, M., Nakamura, M., Ninomiya, N., Nishiyori, H., Nomura, K., Ohno, M., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shiraki, T., Tagami, M., Tagami, Y., Waki, K., Watahiki, A., Muramatsu, M. and Hayashizaki, Y.
	8	Direct Submission Submitted (30-MAR-2004) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa, 230-0045, Japan (E-mail: genome-res@gs.riken.jp, URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)
	8	CDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. Please visit our web site for further details. URL: http://genome.gsc.riken.jp/ Location/Qualifiers 1. 2469 /organism="Mus musculus" /mol_type="mRNA" /db_xref="FANTOM DB:G830026003" /db_xref="taxon:10090" /clone="G830026003"
TITLE JOURNAL PUBMED REFERENCE AUTHORS	COMMENT	
	FEATURES	
	source	





Crowe, M.L., Dalla, E., Dalrymple, B.P., de Bono, B., Della Gatta, G., di Bernardo, D., Down, T., Engstrom, P., Fagioli, M., Faulkner, G., Fletcher, C.F., Fukushima, T., Furuno, M., Futaki, S., Gariboldi, M., Georgii-Hemming, P., Gingras, T.R., Gojibori, T., Green, R.E., Gustincich, S., Harbers, M., Hayashi, Y., Hensch, T.K., Hirokawa, N., Hill, D., Huminicki, L., Iacono, M., Ikeo, K., Iwama, A., Ishikawa, T., Jakt, M., Kanapin, A., Kato, H., Kawasawa, Y., Kelsa, J., Kitamura, H., Kitano, H., Kollias, G., Krishnan, S.P., Kruger, A., Kummerfeld, S.K., Kurochkin, I.V., Lareau, L.P., Lazarevic, D., Lipovich, L., Liu, J., Liuni, S., McWilliam, S., Madan Babu, M., Mader, M., Marchionni, L., Matsuda, H., Matsuzawa, S., Miki, H., Mignone, F., Miyake, S., Morris, K., Mottagui-Tabar, S., Mulder, N., Nakano, N., Nakauchi, H., Ng, P., Nilsson, R., Nishiguchi, S., Nishikawa, S., Nori, F., Ohara, O., Okazaki, Y., Orlando, V., Pang, K.C., Pavan, W.J., Pavese, G., Pesole, G., Petrovsky, N., Piazza, S., Reed, J., Reid, J.F., Ring, B.Z., Ringwald, M., Rost, B., Ruan, Y., Salzberg, S.L., Sandelin, A., Schneider, C., Schonbach, C., Sekiguchi, K., Semple, C.A., Seno, S., Sessa, L., Sheng, Y., Shibata, Y., Shimada, H., Shinada, K., Silva, D., Sinclair, B., Sperling, S., Stupka, E., Sugura, K., Sultana, R., Takenaka, Y., Taki, K., Tammoja, K., Tan, S.L., Tang, S., Taylor, M.S., Tegner, J., Teichmann, S.A., Ueda, H.R., van Nimwegen, E., Verardo, R., Wei, C.L., Yagi, K., Yamaniishi, H., Zabarovsky, E., Zhu, S., Zimmer, A., Hilde, W., Bult, C., Grilmond, S.M., Teasdale, R.D., Liu, E.F., Bruscia, V., Quackenbush, J., Wahlestedt, C., Mattick, J.S., Hume, D.A., Kai, C., Sasaki, D., Tomaru, Y., Fukuda, S., Kanamori-Katayama, M., Suzuki, M., Aoki, J., Arakawa, T., Iida, J., Imamura, K., Itoh, M., Kato, T., Kawaji, H., Kawagashira, N., Kawashima, T., Kojima, M., Kondo, S., Konno, H., Nakano, K., Ninomiya, N., Nishio, T., Okada, M., Plessic, C., Shibata, K., Shiraki, T., Suzuki, S., Tagami, M., Waki, K., Watahiki, A., Okamura-Oho, Y., Suzuki, H., Kawai, J. and Hayashizaki, Y. PANTUM Consortium

The transcriptional landscape of the mammalian genome  
Science 309 (5740), 1559-1563 (2005)  
16141072

7  
Katayama, S., Tomaru, Y., Kasukawa, T., Waki, K., Nakanishi, M., Nakamura, M., Nishida, H., Yap, C.C., Suzuki, M., Kawai, J., Suzuki, H., Carninci, P., Hayashizaki, Y., Wells, C., Frith, M., Ravasi, T., Pang, K.C., Hallinan, J., Mattick, J., Hume, D.A., Lipovich, L., Batalov, S., Engstrom, P.G., Mizuno, Y., Faghihi, M.A., Sandelin, A., Chalk, A.M., Mottagui-Tabar, S., Liang, Z., Lenhard, B. and Wahlestedt, C.

8 (bases 1 to 2430)  
Arakawa, T., Carninci, P., Fukuda, S., Hashizume, W., Hayashida, K., Hori, F., Iida, J., Imamura, K., Inotani, K., Itoh, M., Kanagawa, S., Kawai, J., Kojima, M., Konno, H., Murata, M., Nakamura, M., Ninomiya, N., Nishiyori, H., Nomura, K., Ohno, M., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shiraki, T., Tagami, M., Tagami, Y., Waki, K., Watahiki, A., Muramatsu, M. and Hayashizaki, Y.

Direct Submission  
Submitted (30-MAR-2004) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa, 230-0045, Japan (E-mail: genome-res@sc.riken.jp, URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.  
URL: http://genome.gsc.riken.jp/  
Please visit our web site for further details.  
URL: http://fantom.gsc.riken.jp/  
Location/Qualifiers  
1. .2430  
/organism="Mus musculus"  
/mol\_type="mRNA"

FEATURES  
source

/strain="CS7BL/6J"  
/db\_xref="FANTOM DB:IOC0001L06"  
/db\_xref="taxon:10090"  
/clone="IOC0001L06"  
/cell\_type="4 cells"  
/clone\_lib="RIKEN full-length enriched mouse cDNA library"  
/dev\_stages="4 cells embryo"  
104. .2035  
/note="unnamed protein product; maternal embryonic leucine zipper kinase (MGL) [MGI:106924 GB|NM\_010790, evidence: BLASTN, 100%, match=2428]  
putative"  
/codon\_start=1  
/protein\_id="BAE26362.1"  
/db\_xref="GI:74208327"  
/translation="MKDYDELLKYYELVETIGTGFAKVKLACHVLTGEMVAIKIMDK  
NALGSDLPRVKTETIDALKSRHQHICQLYHVLTKNKLPMVLEYCPGGELPDYIISOD  
RUSEBTRVFRVLIISAVAYVHSQGTAFHRLDKPENLLFDENHKLKILDFGLCAKPKGN  
KDYHLQTCGSLAYAAPELIQKSYLGSEADVMSGILLYVLMCFDPDFDDNDVWALY  
KKIMRGKTEVPKWLSPSILLQQLQVDPKKRISMRNLNHPWQDYSCVQWQSK  
TPLTHLDECVTELSVHRSRQTMEDLISSWOYDHLTATYLLLLAKRGKPARLQL  
LSPSCGTASTTPKSKNLISLEDMSTDNDNCVAGLIDYELCEDKLLAPKTPQVTKHLAES  
NHAASKSPAPGVRRAVANKLMDKENVCTPKSSVKNEQFVSEPKIPVSKNOYKEIIP  
ASPTFPPTPKARAOCLREAPVTPGNSAGADLTITGIVISPPERRCRSMVDVLAQHME  
DTPKKKGTNVFQSLERGLDKVLITALTNRNKKGSARDGPRKRLKHYNTTIRLVNPDQL

Query Match 62.5%; Score 1543.4; DB 6; Length 2430;  
Best Local Similarity 80.5%; Pred. No. 0;  
Matches 1906; Conservative 0; Mismatches 421; Indels 40; Gaps 7;

QY 112 GCGTGGCCCTCGCCCTCAGGTTCTTTTCTTAATTCCAAATAAACTTGCAGAGAGACTA 171  
DB 45 GCTGGGCCCGCTCGCTCAGGTTCTTGTGTGATTAATTAACCTTCCAGGTGACTA 104

QY 172 TGAAGAGATTATGATGAACCTTCTCAATATATGAATATACATGAACATATTGGCAGAGTG 231  
DB 105 TGAAGAGATTATGACGAACCTCTCAATATATGAATATATGAACATATTGAACAGATTGGGACAGGTG 164

QY 232 GCTTTGCAAGGTCAAACTTGGCTGCATATCTTCTACTCGAGAGATGGTAGCTATAAAAA 291  
DB 165 GCTTTGCAAGGTCAAACTGGGCTGCATATCTTCTACTCGAGAGATGGTAGCTATAAAAA 224

QY 292 TCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 351  
DB 225 TCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 284

QY 352 TGAAGAGATTATGATGAACCTTCTCAATATATGAATATACATGAACATATTGGCAGAGTG 411  
DB 285 TGAAGAGATTATGACGAACCTCTCAATATATGAATATATGAACATATTGAACAGATTGGGACAGGTG 344

QY 412 AAATATTCATGTTCTTGAAGTACTGCTCGCTGAGGAGAGCTGTTTGAAGTACTATAATTTCCC 471  
DB 345 AAATATTCATGTTCTTGAAGTACTGCTCGCTGAGGAGAGCTGTTTGAAGTACTATAATTTCCC 404

QY 472 AGGATCGCTGTGAG 531  
DB 405 AGGATCGCTGTGAG 464

QY 532 CTTATGTGACAGCCAGGCTATGTCTCAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 591  
DB 465 CGTATGTGACAGCCAGGCTATGTCTCAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 524

QY 592 ATGAATATCATAAATTAAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 651  
DB 525 ATGAATATCATAAATTAAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 584

QY 652 AGGATTCATCTACAGACATCTGTGGAGTCTGGCTTATGAGCAGCCTGAGTCTGAGTTAATAC 711  
DB 585 AGGATTCATCTACAGACATCTGTGGAGTCTGGCTTATGAGCAGCCTGAGTCTGAGTTAATAC 644

QY 712 AAGGCAATCATATCTTGGATCAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 771  
DB 645 AAGGCAATCATATCTTGGATCAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 704

QY	772	TTCTTATGTGTGGATTTCTACATTTTGAATGATGATAAATGTAAATGGCTTTTATACAGAAGA	831
DB	705	TGCTCATGTGTGTGGATTTCTACCATTTTGAATGATGATAAATGTATGGCTTTGTACAGAAGA	764
QY	832	TTATCAGAGGAAAAATATGATGTGTCCCAAGTGGCTCTCTCCAGATAGCAATCTCGTCTTCTTC	891
DB	765	TAATGAGAGGAAAAATACGAAGTTCTTAAGTGGCTCTCTCCAGATAGCAATCTCGTCTTCTTC	824
QY	892	AACAAATGCTGCAGGTGGACCCAAAGAAACGGATTTCTATGAAAAATCTATGAAACCATC	951
DB	825	AGCAGATGTTGCAGGTGGACCCAAAGAAACGGATTTCTATGAGAAATCTCTCTGAAACCATC	884
QY	952	CCTGGATCATGCAAGATTACAACTATCTCTGTGTAGTGCAAGCAAGAAATCCTTTTATTC	1011
DB	885	CCTGGGTCAATGCAAGATTACAGCTGTCTCCGTGGAGTGGCAAGCAAGAACTCCTTTGATCT	944
QY	1012	ACCTCGATGATGATTCGCTTAACGAACTTTCTGTATCATCAGAAACAAACAGGCAAAACAA	1071
DB	945	ACCTCGATGAGGATTCGCTGAACAGAGCTTTCTGTATCATCACCGCAGCAGGAGGAGCAAA	1004
QY	1072	TGGAGGATTTAATTTCACTGTGGCAGTAGTATCACTCACCTCACGGCTACCTATCTCTGTCTTC	1131
DB	1005	TGAGGATTTAATTTCTGTGTGGCAGTAGTATCACTCACAGCCACCTACCTCTCTGTCTTC	1064
QY	1132	TAGCCAAAGGCTCGGGGAAAAACAGTTCGTTTAAGGCTTTCTCTTTCTCTCTGTGGAC	1191
DB	1065	TAGCCAAAGGCCCGGGGAAAGCGGCTGCTGTACAGCTCTCTGTCTCTCTCTGTGGAA	1124
QY	1192	AAGCCAGTCTACCCCATTCACAGACATCAAGTCAAAATTAATTTGGAGTCTGGAAGATGGA	1251
DB	1125	CGCGCAGCACCACTCCA-----AAGTCAAGAAATCTGAGCCTGGAAAGATATGA	1172
QY	1252	CGCAAGTGATAAAAATTTATGTGGCGGATTAATAGACTATGATTGGTGTGAAGATGATT	1311
DB	1173	GCACAAGTGATGATTAATCTGTGTGGCTGGATTTGATAGACTATGAATTTGTGTGAAGATTAAT	1232
QY	1312	TATCAACAGGTCGTCTACTCCCGGAACATCACACAGTTTACCAGTACTTGGACAGAATCAA	1371
DB	1233	TATTA-----GCTCCCAAGACGCCACAGGTTTACAAACACTTGGCAGAAATCAA	1260
QY	1372	ATGGGTGGAAATCTAAATCATTTAACTCAGCGCTTATGCAGAAACCTGCGAAATAAATTA	1431
DB	1281	ATCAGCGCAGCATCTAAATCACAGCGCCAGCGGATACGCAGACAGTGTGCAAAATAAATTA	1340
QY	1432	AGACAAAGAAATGTATATACTCTCTAGTCTGTGTAAAGAAATGAAGAGTACTTTATGT	1491
DB	1341	TGGACAAAGAAATGTGTGCACTCCCAAGTCTCTGTGAAGAAATGAAGACGAGTTTGTAT	1400
QY	1492	TTCTGTAGCCCAAAGACTCCAGTTAATAGAACCAAGCATTAAGAGAGAAATACTCTACTACGC	1551
DB	1401	TTTCTGAGCCGAGATTTCCAGTTAGTAAGAACCAAGTATAGAGAGAAATACCCGCGCTCAC	1460
QY	1552	CAAAATCGTTTACATACACCTCAAAGCTAGAAACCAAGTGCCTGAAAGAAATCTCAATTA	1611
DB	1461	CAACCCGTTTCCCAACACTGCAAAAGCTAGAGCCAGTGTGCTGTGAGAGAGAGCCCGGTTA	1520
QY	1612	AAATACCAAGTAAATCAACAGGAACAGACAGTTTAATACACAGGTGTCTATAGCCCTGAGA	1671
DB	1521	GAAACACCAAGGAAATCCGACAGGAGCAGACACACTAAGCAGAGGTGTCTATAGCCCCGAGA	1580
QY	1672	GGCGTGCCTCAGTGGAAATTTGGATCTCAACCAAGCACATATGGAGGAGACTCCAAAAA	1731
DB	1581	GGAGGTGCGGTTCAATGSACTGGATCTCAACCAAGGCACATGTGAGGAGTATACCCGAAA	1640
QY	1732	GAAAGGGACCAAGTGTTCGGAGCCTTGAAAGGGGGTGGATAGGTTTATCACTGTGC	1791
DB	1641	AGAAAGGAACCAATGTGTTCGGAGCCTTGAGAGAGGACTGGAATAAGTTCTCTACTCGC	1700
QY	1792	TCACGAGGACCAAGAGGAGGGTTCTGCCAGAGACGGGCCGAGAAAGTCTAAGCTTCACT	1851
DB	1701	TCACAGGAAACAAAGAAAGGGCTCTGCGACAGATGGACCAAGAAAGCGAAAGCTGCACT	1760

QY	1952	ATAATGTGAC	CTACAACTAGATTAGTGAATCCAGATCAACTGTGTGAATGAATAATATGTCTTA	191
DB	1761	ACAAATGTGACTACAACTCGCCTGGTGGAAACCCCGACCACTCTCTGAGGGAATCATATGGCTA	1820	
QY	1912	TTCTTTCCAAAGAAGCATGTTGACATTTCTGACAAAAGGGTTATACACTCGAAGTGTCAAAACAC	1971	
DB	1821	TTCTTTCCAAAGAGACGTGACATCTCGTACAGAAGGTTACACTCTTAAGTGTCAAAAGC	1880	
QY	1972	AGTCAGATTTTGGGAAAGTGCAATGCAAATTTTGAATTTAGAAGTGTGCCAGCTTCCAAAAC	2031	
DB	1981	AGTCTGATTTTGGCAAAGTGACATGCAGTTTGAATCTGGAAGTGTGCCAGCTGCAGAGAC	1940	
QY	2032	CCGATGTGGTGGTATCAGGAGGCGCGCTTAAGGGCGAGTCCCTGGGTTTACAAAAGAT	2091	
DB	1941	CTGACGTGTAGGATCCGGAGACAGCGGCTGAAGGGTGTATGCTCTGGGTTTACAAAGAGAT	2000	
QY	2092	TAGTGGGAAGACATCCTATCTAGCTGCAAGGTATTAATTTGATGGATTCTTCCATCCTGCCGG	2151	
DB	2001	TAGTGGGAAGATATCTTGTCTGGCTGCAAGATGTGACCGATGACAGTGTCTGCTGCGCCGG	2060	
QY	2152	ATGAGTGTGGTGTGATACAGCCTACATAAAGACTGTATTGATCGCTTTGATT-----T	2205	
DB	2061	ATGAATGTGGTGTGATGTTGCAACACAGGAAGACTGTGAGATTGGTTCCATTGGGGGCC	2120	
QY	2206	TAAAGTTCATTGGAACTACCAACTGTGTTCTTAAAGAGCTATCTTAAAGCAATATCTCTT	2265	
DB	2121	ATTGGATCTATGAACTACAAAGCTCATTTAAATAGCCATCTTCAAGACCAATGTATTTTT	2180	
QY	2266	TGTTTTTAAACAAAGATATATTTTGTGTATGAACTCTTAATCTCAAGCCCATCTGTCAATTA	2325	
DB	2181	GCTTTTTTAAACAAAAATTTTACTTTTGGCTGAATCCAAAGGCAA-CTGTCTGTGTCAATTA	2219	
QY	2326	T--GTACTGTCTTTTTTAAATCATGTGGTTTTGTATATTAATAATCTTGACTTTCTTAG	2383	
DB	2240	TCAATTACTCTCTGTTTTTAATCATG--GGCTTGGTATGTAGTAACG-----TTTTTAT	2292	
QY	2384	ATTCACCTTCATATGTGAATGTAAAGTCTTAACTATGCTCTTTTGTAAATGTGTAATTTCT	2443	
DB	2293	ATTCACCTTCCATCAGGGGATGCCAGCTCTTCACTGTGACTCATTTTGTATGTACAGTTTCT	2352	
QY	2444	TTCTGAAATAAACCATTGTGAAATAT	2470	
DB	2353	TTCTGAACTAAACCATTGTGAAATAT	2379	
RESULT 4				
AK011932		2914 bp	linear	HTC 02-SEP-2005
LOCUS	AK011932	2914 bp	linear	
DEFINITION	Mus musculus 10 days embryo whole body cDNA, RIKEN full-length enriched library, clone:261029122 product:maternal embryonic leucine zipper kinase, full insert sequence.			
ACCESSION	AK011932			
VERSION	AK011932.1	GI:12848359		
KEYWORDS	HTC; CAP trapper.			
SOURCE	Mus musculus (house mouse)			
ORGANISM	Mus musculus			
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muroidea; Muridae; Murinae; Mus.			
AUTHORS	Carninci,P. and Hayashizaki,Y.			
TITLE	High-efficiency full-length cDNA cloning			
JOURNAL	Meth. Enzymol. 303, 19-44 (1999)			
PUBMED	10349636			
REFERENCE				
AUTHORS	Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.			
TITLE	Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes			
JOURNAL	Genome Res. 10 (10), 1617-1630 (2000)			
PUBMED	11042159			
REFERENCE				
AUTHORS	Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P.,			



Konno,H., Akiyama,J., Nishi,K., Kitsunai,T., Tashiro,H., Itoh,M., Sumi,N., Iehii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A., Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K., Fujiwaka,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.  
RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer  
Genome Res. 10 (11), 1757-1771 (2000)  
11076861

4  
The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.  
Functional annotation of a full-length mouse cDNA collection  
Nature 409, 685-690 (2001)

5  
The FANTOM Consortium, the RIKEN Genome Exploration Research Group Phase I and II Team.  
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs  
Nature 420, 563-573 (2002)

6  
RIKEN Genome Exploration Research Group, Genome Science Group (Genome Network Core Team) and the FANTOM Consortium.  
Antisense Transcription in the Mammalian Transcriptome  
Science 309, 1564-1566 (2005)

7  
The FANTOM Consortium, Riken Genome Exploration Research Group and Genome Science Group (Genome Network Project Core Group).  
The Transcriptional Landscape of the Mammalian Genome  
Science 309, 1559-1563 (2005)

8 (bases 1 to 2914)  
Adachi,J., Aizawa,K., Akahira,S., Akimura,T., Azai,A., Aono,H., Arakawa,T., Bono,H., Carninci,P., Fukuda,S., Fukunishi,Y., Furuno,M., Hanagaki,T., Hara,A., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hori,F., Imotani,K., Iehii,Y., Itoh,M., Izawa,M., Kasukawa,T., Kato,J., Kawai,J., Kojima,Y., Konno,H., Kouda,M., Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Okazaki,Y., Okido,T., Owa,C., Saito,H., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki,D., Shibata,K., Shibata,Y., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F., Tanaka,T., Tejima,Y., Toya,T., Yamamura,T., Yasunishi,A., Yoshida,K., Yoshino,M., Muramatsu,M. and Hayashizaki,Y.  
Direct Submission  
Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa, 230-0045, Japan (E-mail:genome-res@gsr.riken.jp, URL:http://genome.gsc.riken.jp/, Tel:81-45-503-9222, Fax:81-45-503-9216)  
cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.  
Please visit our web site for further details.  
URL:http://genome.gsc.riken.jp/  
URL:http://fantom.gsc.riken.jp/.  
Location/Qualifiers  
1. .2914  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="C57BL/6J"  
/db\_xref="FANTOM DB:2610209122"  
/db\_xref="taxon:10090"  
/clone="2610209122"  
/tissue type="whole body"  
/clone\_lib="RIKEN full-length enriched mouse cDNA library"  
/dev\_stage="10 days embryo"  
195..2126  
/note="unnamed protein product; maternal embryonic leucine

zipper kinase (MGD|MGI:106924)  
putative"  
/codon\_start=1  
/db\_xref="GI:12848360"  
/translation="MKDYDELLKYYELYETIGTGGFAVKLACHVLTCGEMVAIKIMDK  
NALGSDLRVTEIDALKSLRHQHCQLYVLETKNKLFWLEKPCGGLDFYIISQD  
RLSEETRVFVQILSAVYVHSGYAHRLDKPENLLFDENHKLKLDIFGLCAKPKGN  
KDYHQTCCGSLAAAPBLIOGKSYLGSEADVWSMGILLYVLMCGFLPDDDDNNVALY  
KKIMRGKYEVPKWLSPSSILLQOVLDPKKRISMRNLLNHPWMQDYSCVEVQSK  
TPLHLDDEDCVTELSVHSSRQTMEDLISWQYDHLTATYLLLLAKKARGPKARLQ  
LSFCGTAATTPKSKNLSLEDDSDNCVAGLIDYELCEBKLLAPKTPQVTKHLAES  
NHAASKPAPGVRRAVANKMDKENVCTPKSVKNEEQVFSEPKIPVSKYKKEIP  
ASPTFPTPAKARQCLREAVPTPGNSAGADTLTTGVISPERRCRDMVDNQAHQ  
DTPKKGTNVFGLSLERGLDVLTRNKKGSARDGPRKLVNNTVTLVNPDL  
LSEIMAILPKKNVDPVQKGYTLKCQTQSDFGKVTMQFELEVQCLQRPDVVGIRRLK  
GDAWYKRLVEDILSGCKM"  
2899..2904  
/note="putative"  
2914  
/note="putative"  
polyA\_signal  
polyA\_site  
ORIGIN  
Query Match 62.5%; Score 1543.4; DB 6; Length 2914;  
Best Local Similarity 80.5%; Pred. No. 0;  
Matches 1906; Conservative 0; Mismatches 421; Indels 40; Gaps 7;  
QY 112 GCGTGGCCCTCGCCCTCAGGTTCTTTTCTTAATTCCTCAAAATAAACTTGCAGAGAGACTA 171  
Db 136 GCTGGGCCCTGCGCCCTCAGGTTCTGTTGTCATTTGTAATTAACCTTCCAGGTGGACTA 195  
QY 172 TGAAGAGTTATGATGAACCTTCTCAAAATATATGAATATACATGAACCTATTTGGACAGGTG 231  
Db 196 TGAAGAGTTATGACGAACCTCTCTCAAAATACTATGAACCTATGAACAGTGGACAGGTG 255  
QY 232 GCTTTGCAAGGTCAAACTTGGCTGCCATATCTTACTTGGAGAGAGTGTAGTATATAAAA 291  
Db 256 GCTTTGCAAGGTCAAACTTGGCTGCCATATCTTACTTGGAGAGAGTGTAGTATATAAAA 315  
QY 292 TCATGGAATAAAACACACTAGGAGGTGATTTCCCGGATCAAAACGGAGATTGAGGCCCT 351  
Db 316 TCATGGAATAAGTAATGCGCTAGGAGGTGATTTCCCGGATCAAAACGGAGTGCATGCGC 375  
QY 352 TGAAGAACCTGAGACATCAGCATATATGTCAACTCTACCATGTCTAGAGACAGCCAAACA 411  
Db 376 TGAAGAGTCTGAGACATCAGCACATATGTCTAGCTCTACCATGTCTGAGACAAAGAAC 435  
QY 412 AAATATTCATGTTCTTGTAGTACTCCCTGAGGAGAGCTGTTTGAATATATTTTCCC 471  
Db 436 AAATATTCATGTTCTTGTAGTACTCTCCAGGAGAGAGCTGTTTGAATATATCTCCC 495  
QY 472 AGGATCGCTGTGAGAGAGAGACCGGGTGTCTTCCGTGAGATAGTATCTGCTGTTG 531  
Db 496 AGGATCGCTGTGAGAGAGAGACCGGGTGTCTTCCGTGAGATAGTATCTGCTGAGTTG 555  
QY 532 CTTTATGTGCACAGCCAGGGCTATGCTCAAGGAGAGCTCAAGCCAGAGAAATTTGCTGTTG 591  
Db 556 CGTATGTCCACAGCCAGGGCTATGCCCCACAGGAGAGCTCAAAACAGAGAAATTTATTATT 615  
QY 592 ATGAATATCATAAATTAAGCTGATTTGTTGTTCTGTTGTCGCAAAACCAAGGGTAACA 651  
Db 616 ATGAATAATCATAAAGCTGATTTGTTGTTCTGTTGTCGCAAAACCAAGGGTAACA 675  
QY 652 AGGATTCACATCTACAGACATGCTGTTGAGGTCTGCTTATGTCAGCACCTGAGTTAATAC 711  
Db 676 AGACTACCATCTGACAGACGCTGCTGTTGAGGAGCTTGTCTTATGAGAGTCTCTGAATAC 735  
QY 712 AAGGCAAAATCATATCTTGGATCAGAGGAGAGTGTGTTGGAGAGTGGGATCATCTGTTATATG 771  
Db 736 AAGGGAAGTCGTACCTTGGATCAGAGGAGAGTGTGTTGGAGAGTGGGATCATCTCTCTATG 795  
QY 772 TTCCTATGTGTTGATTTCTACCATTTGATGATGATGATGATGATGATGATGATGATGATG 831





Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A., Kasukawa, T., Saito, R., RIKEN integrated sequence analysis (RISA) system-384 format sequencing pipeline with 384 multicapillary sequencer  
 Genome Res. 10 (11), 1757-1771 (2000)  
 11076861  
 4  
 Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y., Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S., Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamanaka, I., Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R., Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T., Fletschmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H., Kuehl, P., Lewis, S., Matsuo, Y., Nikaide, I., Pesole, G., Quackenbush, J., Schriml, L.M., Staubli, P., Suzuki, R., Tomita, M., Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H., Baldarelli, R., Barsch, G., Blake, J., Boffelli, D., Bojunga, N., Carninci, P., de Bona, M.F., Brownstein, M.J., Bult, C., Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D., Hofmann, M., Hume, D.A., Kamiya, M., Lee, N.H., Lyons, P., Marchionni, L., Mashima, J., Mazzarelli, J., Mombaerts, P., Nordone, P., Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H., Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K.F., Suzuki, H., Toyooka, K., Wang, K.H., Weitz, C., Whittaker, C., Wilming, L., Wyshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawai, J., Kohetsuki, S., and Hayashizaki, Y.  
 RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium  
 Functional annotation of a full-length mouse cDNA collection  
 Nature 409 (6821), 685-690 (2001)  
 11217851  
 5  
 Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S., Nikaide, I., Osato, N., Saito, R., Suzuki, H., Yamanaka, I., Kiyosawa, H., Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A., Schonbach, C., Gojobori, T., Baldarelli, R., Hill, D.P., Bult, C., Hume, D.A., Quackenbush, J., Schriml, L.M., Kanapin, A., Matsuda, H., Batalov, S., Beisel, K.W., Blake, J.A., Brad, D., Brusic, V., Chothia, C., Corbani, L.E., Cousins, S., Dalla, E., Dragani, T.A., Fletcher, C.P., Forrest, A., Frazer, K.S., Gaasterland, T., Gariboldi, M., Gissi, C., Godzik, A., Gough, J., Grimmond, S., Gustincich, S., Hirokawa, N., Jackson, I.J., Jarvis, E.D., Kanai, A., Kawai, H., Kwasawa, Y., Kedzierzki, R.M., King, B.L., Konagawa, A., Kurachkin, I.V., Lee, Y., Lenhard, B., Lyons, P.A., Maglott, D.R., Maltais, L., Marchionni, L., McKenzie, L., Miki, H., Negashima, T., Numata, K., Okido, T., Pavan, W.J., Pertea, G., Pesole, G., Petrovsky, N., Pillai, R., Pontius, J.U., Qi, D., Ramachandran, S., Ravasi, T., Reed, J.C., Reed, D.J., Reid, J., Ring, B.Z., Ringwald, M., Sandelin, A., Schneider, C., Sempole, C.A., Setou, M., Shimada, K., Sultana, R., Takenaka, Y., Taylor, M.S., Teasdale, R.D., Tomita, M., Verardo, R., Wagner, L., Wahlstedt, C., Wang, Y., Watanabe, Y., Wells, C., Wilming, L.G., Wynshaw-Boris, A., Yanagisawa, M., Yang, I., Yang, L., Yuan, Z., Zavalan, M., Zhu, Y., Zimmer, A., Carninci, P., Hayatsu, N., Hirozane-Kishikawa, T., Konno, H., Nakamura, M., Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, J., Aizawa, K., Arakawa, T., Fukuda, S., Hara, A., Hashizume, W., Imotani, K., Ishii, Y., Itoh, M., Kigawa, I., Miyazaki, A., Sakai, K., Sasaki, D., Shibata, K., Shinagawa, A., Yasunishi, A., Yoshino, M., Waterston, R., Lander, E.S., Rogers, J., Birney, E. and Hayashizaki, Y.  
 FANTOM Consortium  
 Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs  
 Nature 420 (6915), 563-573 (2002)  
 12466851  
 6  
 Carninci, P., Kasukawa, T., Katayama, S., Gough, J., Frith, M.C., Maeda, N., Oyama, R., Ravasi, T., Lenhard, B., Wells, C., Kodzius, R., Shimokawa, K., Bajic, V.B., Brenner, S.E., Batalov, S., Forrest, A.R., Zavolan, M., Davis, M.J., Wilming, L.G., Aidinis, V., Allen, J.E., Ambesi-Implombato, A., Apweiler, R., Aturaliya, R.N., Bailey, T.L., Bansal, M., Baxter, L., Beisel, K.W., Bersano, T., Bono, H., Chalk, A.M., Chiu, K.P., Choudhary, V., Christoffels, A., Clutterbuck, D.R.,

Crowe, M.L., Dalla, E., Dalrymple, B.P., de Bono, B., Della Gatta, G., di Bernardo, D., Down, T., Engstrom, P., Fagioli, M., Faulkner, G., Fletcher, C.F., Fukushima, T., Furuno, M., Futaki, S., Gariboldi, M., Georgii-Hemming, P., Gingeras, T.R., Gojobori, T., Green, R.E., Gustincich, S., Harbers, M., Hayashi, Y., Hensch, T.K., Hirokawa, N., Hill, D., Hummel, L., Iacono, M., Ikeo, K., Iwama, A., Ishikawa, T., Jakt, M., Kanapin, A., Katoh, M., Kawasawa, Y., Kelsso, J., Kitamura, H., Kitano, H., Kollias, G., Krishnan, S.P., Kruger, A., Kummerfeld, S.K., Kurochkin, I.V., Lareau, L.P., Lazarevic, D., Lipovich, L., Liu, J., Liuni, S., McWilliam, S., Madan, Babu, M., Madera, M., Marchionni, L., Matsuda, H., Matsuzawa, S., Miki, H., Mignone, P., Miyake, S., Morris, K., Mottagui-Tabar, S., Mulder, N., Nakano, N., Nakaguchi, H., Ng, P., Nilsson, R., Nishiguchi, S., Nishikawa, S., Nori, F., Ohara, O., Okazaki, Y., Orlando, V., Pang, K.C., Pavan, W.J., Pavese, J., Pesole, G., Petrovsky, N., Piazza, S., Reed, J., Reid, J.F., Ring, B.Z., Ringwald, M., Rost, B., Ruan, Y., Salzberg, S.L., Sandelin, A., Schneider, C., Schonbach, C., Sekiguchi, K., Sempole, C.A., Seno, S., Sessa, L., Sheng, Y., Shibata, Y., Shimada, H., Shimada, K., Silva, D., Sinclair, B., Sperling, S., Stupka, S., Sugita, K., Sultana, R., Takenaka, Y., Taki, K., Tammoja, K., Tan, S.L., Tang, S., Taylor, M.S., Tegner, J., Teichmann, S.A., Ueda, H.R., van Nimwegen, E., Verardo, R., Wei, C.L., Yagi, K., Yamanishi, H., Zabarovsky, E., Zhu, S., Zimmer, A., Hide, W., Bult, C., Grimmond, S.M., Teasdale, R.D., Liu, E.T., Brusic, V., Quackenbush, J., Wahlstedt, C., Mattick, J.S., Hume, D.A., Kai, C., Sasaki, D., Tomaru, Y., Fukuda, S., Kanamori-Katayama, M., Suzuki, M., Aoki, J., Arakawa, T., Iida, J., Imamura, K., Itoh, M., Kato, T., Kawaji, H., Kawagashira, N., Kawashima, T., Kojima, M., Kondo, S., Konno, H., Nakano, K., Ninomiya, N., Nishio, T., Okada, M., Plessey, C., Shibata, K., Shiraki, T., Suzuki, S., Tagami, M., Waki, K., Watahiki, A., Okamura-Oho, Y., Suzuki, H., Kawai, J. and Hayashizaki, Y.  
 FANTOM Consortium  
 The transcriptional landscape of the mammalian genome  
 Science 309 (5740), 1559-1563 (2005)  
 16141072  
 7  
 Katayama, S., Tomaru, Y., Kasukawa, T., Waki, K., Nakanishi, M., Nakamura, M., Nishida, H., Yagi, C.C., Suzuki, M., Kawai, J., Suzuki, H., Carninci, P., Hayashizaki, Y., Wells, C., Frith, M., Ravasi, T., Pang, K.C., Hallinan, J., Mattick, J., Hume, D.A., Lipovich, L., Batalov, S., Engstrom, P.G., Mizuno, Y., Faghihi, M.A., Sandelin, A., Chalk, A.M., Mottagui-Tabar, S., Liang, Z., Lenhard, B. and Wahlstedt, C.  
 RIKEN Genome Exploration Research Group  
 Antisense transcription in the mammalian transcriptome  
 Science 309 (5740), 1564-1566 (2005)  
 16141073  
 8 (bases 1 to 2447)  
 Arakawa, T., Carninci, P., Fukuda, S., Hashizume, W., Hayashida, K., Hori, F., Iida, J., Imamura, K., Imotani, K., Itoh, M., Kanagawa, S., Kawai, J., Kojima, M., Konno, H., Murata, M., Nakamura, M., Ninomiya, N., Nishiyori, H., Nomura, K., Ohno, M., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shiraki, T., Tagami, M., Tagami, Y., Waki, K., Watahiki, A., Muramatsu, M. and Hayashizaki, Y.  
 Direct Submission  
 Submitted (14-APR-2004) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa, 230-0045, Japan (E-mail: genome-res@sc.riken.jp, URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)  
 cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in RIKEN Genomic Sciences Center and Genome Science Laboratory in RIKEN, Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.  
 Please visit our web site for further details.  
 URL: http://genome.gsc.riken.jp/  
 URL: http://fantom.gsc.riken.jp/  
 Location/Qualifiers  
 1. .2447  
 /organism="Mus musculus"  
 /mol\_type="mRNA"

```

/strain="C57BL/6J"
/db_xref="FANTOM_DB:I530014L15"
/db_xref="taxon:10090"
/clone="I530014L15"
/sex="female"
/tissue_type="placenta"
/clone_lib="Riken full-length enriched mouse cDNA library"
/dev_stage="12 days pregnant adult"
119. .2050
/note="unnamed protein product; maternal embryonic leucine
zipper kinase (MGD|MG1:106924 GB|NM_010790, evidence:
BLASTN, 99%, match=2446)
putative"
/codon_start=1
/protein_id="BAE39542.1"
/db_xref="GI:74195447"
/translation="MKDYDELLKYYELYETIGTGGPAKVKLACHVLTGEMVAIKMDK
NALGSDLPVKTEIDALKSLRHQHCQLYHVLETKNKLFWLVEYCPGGEHFDYIIISOD
RLSEETRVFRQLSVAIVYGYSHARDLPDENHKLKGLIDFLGCAKPGKN
KDYHQLQCCSLAYAPELIGKSYLSHEDVMSMGLIYLMCGFLPFPDDDMVALY
KTMIRGKEVPKMLSPSIIILLQMLQVDPKRIISMRNLNHPWVMDQYDSCPVMWOSK
TPLTHLEDCEVTLSVHRRSSROTMELISSWOYDHLTYALLLLAKKAGKPARLRLL
LSTSCGPASTPTPKSKNLSDMSTSDNCCVAGLIDVELCEDKLLAKPTPOVTKHLAES
NHAASKSPAGVRVANLKDRENVCTPKSSVKNNEQVFPSEPKIPVSKNKKYKREIP
ASPTPTTPAKARAQRCURVAPRTGNSGADTLTTGTVISPERRRCSMDVDLNOAHWE
Query Match 62.4%; Score 1541.8; DB 6; Length 2447;
Best Local Similarity 80.5%; Pred. No. 0;
Matches 1905; Conservative 0; Mismatches 422; Indels 40; Gaps 7;

Qy 112 GCCGTGCCCTCCGCCCTCAGGTTCTTTTCTTAATCCAAATAACTTGCAAGGAGCTA 171
Db 60 GCTGGGCCCGCTGCCCTTCAGGTTCTGGTTGTCAATTTTGAATAACCTTCACAGGTGGACTA 119

Qy 172 TGAAGGATTATGATGAATCTCTCAAAATATATGATTAACATGAACTATTTGGGACGAGTG 231
Db 120 TGAAGGATTATGACGAATCTCTCAAAATATATGATTAACATATGAAACGATTTGGGACGAGTG 179

Qy 232 GCTTTGCAAAAGGTCAAACTTGCCTGCCATATCTCTTACTGCGAGAGATGCTAGCTATAAAAA 291
Db 180 GCTTTGCAAAAGGTCAAACTGCGCTGCCATGCTCTCACTGGAGAGATGGTAGCTATAAAAA 239

Qy 292 TCATGGATAAAAAACACACTAGGAGTGATTTGCCCGCGATCAAAACGGAGATTGAGGCCT 351
Db 240 TCATGGATAAAGATGCGCTAGGAGTGATTTGCCCGGAGTCAAAACCTGAGATCGATGCGC 299

Qy 352 TGAAGAACCTGAGACATCAGCATATATGTCAACTTACCATGTGCTGAGACAGCCCAACA 411
Db 300 TGAAGAGCTCTGAGACATCAGCACATATGTGAGCTCTACCATGTGCTGGAGCAAAAGAAC 359

Qy 412 AAATATTTCATGGTTCTTGAGTACTGCCCTGGAGGAGAGCTGTTTGGACTATATAATTTCCC 471
Db 360 AAATATTTCATGGTTCTGAGTACTGTTCCAGAGGAGAGAGCTGTTTGGACTACATAATCTCCC 419

Qy 472 AGGATCGCCTGTTCAGAAAGAGAGACCCGGGTTGTCTTCCTCGATAGTATCTGCTGTTG 531
Db 420 AGGATCGCCTGTTCGAAAGAGAGAGACCCGGGTGCTGTTCCGTTCAGATACTGCTGCGAGTTG 479

Qy 532 CTTATGTGCAAGCCAGGGCTATGCTCAAGGGAGACCTCAAGCCAGAAAAATTTGCTGTTTG 591
Db 480 CGTATGTGCCACAGCCAGGGCTATGCCCAAGGGACCTCAAAACCCAGAAAAATTTTATTATTG 539

Qy 592 ATGAATATCATAAATTAAGCTGATGACATTTGCTCTCTGTCGAAAACCCAGGGTAACA 651
Db 540 ATGAATAATCATAAGCTAAAGCTGATGACATTTGCTCTTTGTGTAACCCAGGGCAACA 599

Qy 652 AGGATTACCATCTACAGACATGCTGTGGGAGTCTGGCTTATGCAAGCACCTTGAGTTAATAC 711
Db 600 AGGACTACCATCTCAGACCGTGTGTGGAGCGCTTGCTTATGAGCTCTCTGAACATAATC 659

Qy 712 AAGCAAAATCATATCTTTGGATTCAGAGCAGATGTTTGGAGCATGGGGCATACTGTTATATG 771
Db 660 AAGGGAAGTCGTACCTTGGATTCAGAGCAGATGTTTGGAGCATGGGGCATCTCTCTGTATG 719

```

Qy	772	TTCTTATGCTGTGGATTTCTACCAATTTGATGATGATAATGTAAATGGCTTTTATACAAAGA	831
Db	720	TGCTCATGTGTGGATTTTCTACCAATTTGATGATGATAATGTTCATGGCTTTGTACAAAGA	779
Qy	832	TTATGACAGAGAAATATGATGTTTCCCAAGTGGCTCTCTCCAGTAGCATTTCTGCTTCTTC	891
Db	780	TAATGAGAGGGAATACGAATTTCTTAAGTGGCTCTCTCCAGTAGCATTTCTGCTTCTCC	839
Qy	892	AACAAATGCTGCGAGTGGACCCAAAGAAACGGATTTTCTATGAAATAATCTATTTGAACCATC	951
Db	840	AGCAGATGTTTGCAGGTGGACCCAAAGAAACGGATTTTCTATGAGAAATCTCTCTGAACCATC	899
Qy	952	CCTGGATCATGCAAGATTACAACTATCTCTGTGAGTGGCAAGCAAGATCCCTTTTATTC	1011
Db	900	CCTGGGTCTATGCAAGATTACAGCTGTCCCGTGGAGTGGCAAGCAAGATCCCTTTGACTC	959
Qy	1012	ACCTCGATGATGATGCGTAAACAGAACTTTCTGTATACATCACAGAAACAAACAGGCAACAA	1071
Db	960	ACCTCGATGAGGATTTGCGTGACAGAGCTTCTTGATCATCACCGCAGCAGCAGGACAGACAA	1019
Qy	1072	TGGAGGATTTAAATTTCACTGTGGCAGTATGATCACCTCAGCGCTACCTATCTTCTGCTTC	1131
Db	1020	TGGAGGATTTAAATTTGCTGTGGCAGTACATCACCTCAGCGCACCTACCTTCTGTGCTTC	1079
Qy	1132	TAGCCAAGAGGCTTCGGGGAACCAAGTTCGTTTAAAGGCTTTCTTCTTCTCTCTGTGGAC	1191
Db	1080	TAGCCAAGAGGCCCGGGGAGCGGCTGCTACGGCTCCTGCTCTCTTGTGGAA	1139
Qy	1192	AAGCAGTGCTACCCCATTCACAGACATCAAGTCAATAATTTGGAGTCTGGAGAGATGTA	1251
Db	1140	CCGCCAGCACCATCCA-----AAGTCAAAGAACTCTGAGCGCTGGAGAGATATGA	1187
Qy	1252	CCGCAAGTATAAAAAATTATGTGGCGGATTAATAGACTATGATTTGGTGTGAAGATGATT	1311
Db	1188	GCACAGTGATGATTACTGTGTGCTTGGATTTGATAGCTATGAAATTTGTGTGAAGATTAAT	1247
Qy	1312	TATCAACAGGTGCTGTACTCCCGCAACATCACAGTTTACCAAGTACTGGACAGAAATCAA	1371
Db	1248	TAATTA-----GCTCCCAAGAGGCCACAGGTTACCAAACTTGGCAGAAATCAA	1295
Qy	1372	ATGGGTGGAAATCTAAATCATTAATCTCAGCCTTATGCAGAAACACCTGGCAATTAATTA	1431
Db	1296	ATCACGCGAGCATCTAAATCACACGCGCCAGGGGTACGACAGCAGTGGCAAAATAAATTA	1355
Qy	1432	AGAACAAAGAAAATGTATATCTCTCTAAAGTCTGCTGTAAAGAAATGAAGAGTACTTTATGT	1491
Db	1356	TGACAAAGAAATGTGTGCACTCCCAAGTCTTCTGTGAAGATGAAGACAGTTTGTAT	1415
Qy	1492	TTCCCTGAGCCAAAGACTCCAGTTTAATAAGAAACAGCATTAAGAGAGAAATACTCACTACGC	1551
Db	1416	TTTTCTGAGCCGAAGATTCAGTTTAGTAAAGAAACAAGTATAAGAGAGAGAAATACCGCGCTCAC	1475
Qy	1552	CAAAATCGTTACACTACACCTCTAAAAGCTAGAAACAGTGCCTCTGAAAGAAATCTCCAATTA	1611
Db	1476	CAACCCGTTTCCCAACACTGTCAAAAGCTATAGCCCAAGTGCCTGTAGAGAAAGCCCCGGTTA	1535
Qy	1612	AAATACAGTAAATTCACAGGAACACAGAAAGTTAATGACAGGTGTCAATTAGCCCTGAGA	1671
Db	1536	GAACACACCGGAATTCGCGAGGAGCAGACACACTAACGACAGGTGTCAATTAGCCCCGAGA	1595
Qy	1672	GGCGGTGCCCTCAGTGGAAATTTGGATCTCAACCAACACATATGGAGGAGACTCCCAAAA	1731
Db	1596	GGAGGTGCCGTTCAATTTGGACGTGGATCTCAACACAGGCACACTGGAGGATACCCCCGAAA	1655
Qy	1732	GAAGAGGAGCCAAAGTGTTTTGGGAGCCTTGAAAGGGGGTTGGATAAGGTTTATCACTGTGC	1791
Db	1656	AGAAAGGAACCAATGTGTTTGGAGCCTTGAGAGAGGACTGGATATAGTTTCTCACTGCCG	1715
Qy	1792	TCAACAGGAGCAAAAGGAGGTTTTCGACAGACAGCGGCCACAGAACTTAAGCTTCACT	1851
Db	1716	TCAACAGGAACAAAGAAAGGGCTCTGCCAGATATGGACCAAGAAACGAAAGCTGCATCT	1775

QY	1852	ATAATGTGACTCAAACTAGATTAAGTGAATCCAGATCAACTGTGTAATGAATAATGTCTTA	1911
Db	1776	ACAATGTGACTCAAACTCGCTCGTGAAACCCGACACAGCTCTCTGAGCGAAATCATGGCTA	1835
QY	1912	TTCTTCCAAAGAAGCATGTTGACTTTTGTACAAAAGGGTTATACACTGAAGTGTCAACAC	1971
Db	1836	TTCTTCCAAAGAAGACGTGGACTTCGTACAGAAAGGTTACACTCTAAAGTGTCAACACG	1895
QY	1972	AGTCAGATTTTGGGAAAGTGCACATGCAAATTTGAAATTAGAAGTGTGCCAGCTTCAAAAAC	2031
Db	1896	AGTCTGATTTTGGCAAGTGCATGACATGAGTTTGNACTGGAAGTGTGCCAGCTCGAGNAC	1955
QY	2032	CCGATGTGTGGGTATCAGGAGGACGCGCTTAAAGGCGATGCTCGGGTTTACAAAAGAT	2091
Db	1956	CTGACGTGGTAGGCATCCGAGACAGCGGCTGAAGGGTGATGCTGGGTTTACAAGAGAT	2015
QY	2092	TAGTGGAGACATCCTATCTAGCTGCAAGGTATTAATTGATGGATTCTTCCATCTGCCGG	2151
Db	2016	TAGTGGAAAGATATCTTGTCTGGCTGCCAAGATGTGACCGGTGCACAGTGTCCGCTGCCCCG	2075
QY	2152	ATGAGTGTGGGTGTGATACAGCCTACATAAAGACTGTTATGATGCTTTTGATT-----T	2205
Db	2076	ATGAATGTGGGTGTGATGTTGCACACAGGAAGACTGTGAGATTGGTTCATTTGGGGGCC	2135
QY	2206	TAAAGTTTCATTGGAACTACCAACTTGTTCCTAAAGAGCTATCTTTAAGACCAATATCTCTT	2265
Db	2136	ATTGGATCTATGGAACCTACAAGCTCAATTTAAATAGCCATCTTCAAGACCAATGATATTTT	2195
QY	2266	TGTTTTTAAACAAAAGATATATTTTGTGTATGNACTTAATCAAGCCATCTGTCAATTA	2325
Db	2196	GCTTTTTTAAACAAAATTTTTTACTTTTGGCTGAATCCAAAGGCAA-CTGTCTCTGTCAATTA	2254
QY	2326	T--GTTACTGTCTTTTTTAATCATGTGCTTTTGTATATTAATTAATCTTGACTTTTCTTAG	2383
Db	2255	TCAATTAATCTCTGTTTTTAATCATG--GGCTTGGTAUGTTAGTAAGTCTG-----TTTTTAT	2307
QY	2384	ATTCACATTCACATATGTGAATGTAGACTCTTAAACTATGCTCTTTTGTAAATGTGAATTTCT	2443
Db	2308	ATTCACATTCACATCAGGGGATGCCAGCTCTTCACCTGTGACTCATTTTGTATGTACAGTTTCT	2367
QY	2444	TTCTGAATATAAACCATTTGTGAATAT	2470
Db	2368	TTCTGAACATAAACCATTTGTGAATAT	2394
RESULT	6		
AKI64138			
LOCUS		2489 bp	mRNA
DEFINITION		Mus musculus 12 days embryo spinal cord cDNA, RIKEN full-length enriched library, clone:CS30047P13 product:maternal embryonic leucine zipper kinase, full insert sequence.	linear HTC 21-SEP-2005
ACCESSION	AKI64138		
VERSION	AKI64138.1	GI:74211111	
KEYWORDS	HTC; CAP trapper.		
SOURCE	Mus musculus (house mouse)		
ORGANISM	Mus musculus		
REFERENCE			
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muroidea; Muridae; Murinae; Mus.		
TITLE	1		
JOURNAL	Carninci, P. and Hayashizaki, Y.		
PUBLISHED	High-efficiency full-length cDNA cloning		
AUTHORS	Math. Enzymol. 303, 19-44 (1999)		
TITLE	10349636		
JOURNAL	2		
PUBLISHED	Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.		
AUTHORS	Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes		
TITLE	Genome Res. 10 (10), 1617-1630 (2000)		
JOURNAL	11042159		
PUBLISHED	3		
AUTHORS	Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,		



QY 712 AAGCAAAATCATATCTTGGATCAGAGCGAGATGTTTGGACATGGGCATCTGTTATATG 771  
DB |||||  
QY 703 AAGGAAGTCGTACCTTGGATCAGAGCAGATGTTTGGAGCATGGGCATCTCTGTATG 762  
DB |||||  
QY 772 TTCTTATGTGTGATTTCTACCAATTTGATGATGATTAATGTAATGGCTTTATACAAAGA 831  
DB |||||  
QY 763 TGCTCATGTGTGATTTCTACCAATTTGATGATGATTAATGTAATGGCTTTGTAAGAAGA 822  
DB |||||  
QY 832 TTATGAGAGGAAAATATGATGTTTCCCAAGTGGCTCTCTCCAGTAGCATTTCTGCTTCTC 891  
DB |||||  
QY 823 TAATGAGAGGAAATACGAAGTTCTTAAGTGGCTCTCTCCAGTAGCATTTCTGCTTCTC 882  
DB |||||  
QY 892 AACAAATGTCAGGTGGAGCCCAAGAAAGCGAATTTCTATGAAAAATCTATTGAACCATC 951  
DB |||||  
QY 883 AGCAGATGTTGAGGTGGAGCCCAAGAAAGAAACGGAATTTCTATGAAATCTCTGAAACCATC 942  
DB |||||  
QY 952 CTTGGATCATGCAAGATTAACATATCTCTGTTGAGTGGCAAGCAAGATCTCTTTATTC 1011  
DB |||||  
QY 943 CTTGGGTCTATGCAAGATTAACAGCTGCTCCGTGGAGTGGCAAGCAAGATCTCTTTGACTC 1002  
DB |||||  
QY 1012 ACCTCGATGATGATTCGTAACAGAACTTTCTGTATCATCACAGAAACAAACAGGCAACAA 1071  
DB |||||  
QY 1003 ACCTCGATGAGGATTCGGTGACAGAGCTTTCTGTATCATCACCGCAGCAGGACGACAA 1062  
DB |||||  
QY 1072 TGGAGGATTTAAATTTCACTGTGGCAGTATGATCACCTCAAGGCTACCTATCTTCTGCTTC 1131  
DB |||||  
QY 1063 TGGAGGATTTAAATTTCTGTGGCAGTACGATCACCTCAAGGCTACCTTCTGCTTC 1122  
DB |||||  
QY 1132 TAGCCAAAGAGGCTCGGGGAAACCGAGTTGTTTAAAGGCTTTCTTCTCTCTGTTGGAC 1191  
DB |||||  
QY 1123 TAGCCAAAGAGGCTCGGGGAAACCGGCTCGTCTACAGCTCTCTTCTTGTGGAA 1182  
DB |||||  
QY 1192 AAGCCAGTGTACCCCATTCACAGACATCAAGTCAAAATTAATTCGAGTCTGGAAGATGTGA 1251  
DB |||||  
QY 1183 CCGCCAGCACCATCCA-----AAGTCAAGAAATCTGAGCTGGGAAGATATGA 1230  
DB |||||  
QY 1252 CGCAAGTGTATAAAATTAATGTGGCGGATTAATAGACTATGATGTTGGTGTGAAGATGAT 1311  
DB |||||  
QY 1231 GCACAAGTGTATGATGATGTTGGCTGGATTTGATAGACTATGAATTTGTGTGAAGATAAT 1290  
DB |||||  
QY 1312 TATCAACAGTGTGCTACTTCCCGGAACATCAGAGTTTACCAAGTACTGGACAGATCAA 1371  
DB |||||  
QY 1291 TATTA-----GCTCCCAAGACGCCACAGGTTACCAACACTTGGCGAGATCAA 1338  
DB |||||  
QY 1372 ATGGGTGGAAATCTAAATCAATTAATCTCAGCCTTATGCAAAACACCTGCATAATTAATTA 1431  
DB |||||  
QY 1339 ATCAGCAGCATCTAAATCACAGCGCCAGGGTACGAGAGCAGTGGCAAAATTAATTA 1398  
DB |||||  
QY 1432 AGAACAAGAAAATGTATATPACTCTTAAGTCTGCTGTAAAGAAATGAAGTACTTTATGT 1491  
DB |||||  
QY 1399 TGGACAAAGAAAATGTGTGCACTCCCAAGTCTTCTGTGAAGAAATGAAGAGCAGTTTGTAT 1458  
DB |||||  
QY 1492 TTCCTGAGCCAAAGACTCCAGTTAATGAACACAGCATTAAGAGAAAATACTACTACGC 1551  
DB |||||  
QY 1459 TTTCTGAGCCGAAGATTTCCAGTTAGTAAGAACACAGTATAGAGAGAAATACCGGCTCAC 1518  
DB |||||  
QY 1552 CAATCGTTTACACTACACCTCAAAAGCTAGAAACAGTGCCTGAAAGAACTCCAATTA 1611  
DB |||||  
QY 1519 CAACCCGTTTCCAAACACTGCAAAAGCTAGAGCCAGTGCCTGAGAGAGCCCGGTTA 1578  
DB |||||  
QY 1612 AAATACCAGTAAATTCACAGGAAACAGACAAAGTTAATGACAGGTGTCAATPAGCCCTGAGA 1671  
DB |||||  
QY 1579 GAACACCGGGAATTCGCAGGAGCAGACACACTAACGACAGGTGTCAATGAGCCCGAGA 1638  
DB |||||  
QY 1672 GCGGTGCGCCTCAGTGGAAATTTGATCTCAACAGCAGATATGAGGAGACTCCAAAAA 1731  
DB |||||  
QY 1639 GGAAGTCCCGTTCAATGGACGTGGATCTCAACAGGACACATGGAAGGATACCCCGGAAA 1698  
DB |||||  
QY 1732 GAAAGGAGGACCAAGTGTTTGGGAGCCTTCAAAAGGGGTTGGATTAAGTTATCACTGTGC 1791  
DB |||||  
QY 1699 AGAAGGAACCAATGTGTTTGGAGGCTTGAAGAGACTGGAATAGGTTCTCACTGCGC 1758  
DB |||||  
QY 1792 TCACAGGAGCAAAAGGAAGGGTTCTGCCAGAGACGCGGCCCAAGAGACTAAAGCTTCACT 1851  
DB |||||

DB 1759 TCACAAGGAACAAGAGAGAGGGCTCTGCCAGAGATGGACCAAGAAAGCGAAAGCTCACT 1818  
QY |||||  
DB 1852 ATAATGTGTAACAATAGATTAGTGAATCCAGATCAACTGTTTGAATGAAATAATGTCTA 1911  
DB |||||  
QY 1819 ACATGTGTAACAATCGCTGCTGGTGAACCCGACCAAGTCTCTGAGCGAAATCATGGCTA 1878  
DB |||||  
QY 1912 TTCTTCCAAAGAGCATGTTGATCTTTGTACAAAAGGGTTATACACTGAAGTGTCAAAACAC 1971  
DB |||||  
QY 1879 TTCTTCCAAAGAAAGACGTGGACTTCGTACAGAAAGGTTTACACTCTTAAAGTGTCAAAACGC 1938  
DB |||||  
QY 1972 AGTCAGATTTTGGGAAAGTGACAATGCAATTTGAATTAGAGTGTGCCAGCTTCAAAAC 2031  
DB |||||  
QY 1939 AGTCTGATTTTGGCAAAAGTGACAATGCAATTTGAATGGAAGTGTGCCAGCTGACAGAC 1998  
DB |||||  
QY 2032 CGATGTGTGTGATATCAGGAGCGAGCGCTTTAAAGGGCGATGCTCGGTTTACAAAAGAT 2091  
DB |||||  
QY 1999 CTGACGTGTGTGATCCGAGACAGCGCTGAGGGTGTGCTGGGTTTACNAGAT 2058  
DB |||||  
QY 2092 TAGTGAAGACATCTCTATCTAGCTGCAAGGTATAATTTGATGGAATCTTCCATCTGCCCG 2151  
DB |||||  
QY 2059 TAGTGAAGATATCTTGTCTGGCTGCAAGATGTGACCGATGACAGTGTCCGTCTGCCCG 2118  
DB |||||  
QY 2152 ATGAGTGTGGTGTGATACAGGCTACATAAAGCTGTATGATCGCTTTGATTT-----T 2205  
DB |||||  
QY 2119 ATGAATGTGGTGTGATGTGCAACAGGAAGACTGTGAGATTTGGTTCCTTTGGGGGCC 2178  
DB |||||  
QY 2206 TAAAGTTCAATTCGAATCTACCACTTGTCTTAAAGAGCTATCTTAAGACCAATATCTCTT 2265  
DB |||||  
QY 2179 ATTGNACTATGGAATCTACAGCTCATTTAAATAGCCATCTTCNAGACCAATGATTTTT 2238  
DB |||||  
QY 2266 TGTTTTAAACAAAAGATATTTTGTGTATGAATCTAAATCTAAATCAAGCCCATCTGTATTA 2325  
DB |||||  
QY 2239 GCTTTTAAACAAAATTTTACTTTGTGGCTGAATCCAAGGCAA-CTGTTCTGTCTATTA 2297  
DB |||||  
QY 2326 T--GTTACTGCTTTTAAATCATGCTGTTTGTATATTAATTAATTTGACTTCTTAG 2383  
DB |||||  
QY 2298 TCAATTTACTCTCTGTTTAAATCATG--GGCTGGTATGTTAGTAACCTG-----TTTTAT 2350  
DB |||||  
QY 2384 ATTCACCTCCATATGGAATGTAAGCTCTTAAGCTATCTCTTGTGTAATGTGTAATTTCT 2443  
DB |||||  
QY 2351 ATTCACCTCCATGAGGATGCGAGCTCTTCACTGTGATCATTTTGTGATGTACAGTTTCT 2410  
DB |||||  
QY 2444 TTCTGAAATPAAACCACTTTGTGAATAT 2470  
DB |||||  
QY 2411 TTCTGAACTAAACCACTTTGTGAATAT 2437  
DB |||||

## RESULT 7

AK161029

LOCUS

DEFINITION

AK161029

VERSION

KEYWORDS

SOURCE

ORGANISM

Mus musculus

Mus musculus

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;

Sciurognathi; Muridae; Murinae; Mus.

1

Carninci, P. and Hayashizaki, Y.

High-efficiency full-length cDNA cloning

Meth. Enzymol. 303, 19-44 (1999)

10349636

2

Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,

Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.

Normalization and subtraction of cap-trapper-selected cDNAs to

prepare full-length cDNA libraries for rapid discovery of new genes

Genome Res. 10 (10), 1617-1630 (2000)

AK161029 2541 bp mRNA linear HTC 21-SEP-2005  
Mus musculus 14 days embryo liver cDNA, RIKEN full-length enriched  
library, clone:4322406P16 product:maternal embryonic leucine zipper  
kinase, full insert sequence.

AK161029.1 GI:74145424

HTC; CAP trapper.

Mus musculus (house mouse)

Mus musculus

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;

Sciurognathi; Muridae; Murinae; Mus.

1

Carninci, P. and Hayashizaki, Y.

High-efficiency full-length cDNA cloning

Meth. Enzymol. 303, 19-44 (1999)

10349636

2

Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,

Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.

Normalization and subtraction of cap-trapper-selected cDNAs to

prepare full-length cDNA libraries for rapid discovery of new genes

Genome Res. 10 (10), 1617-1630 (2000)







URL:http://genome.gsc.riken.jp/ URL:http://fantom.gsc.riken.jp/. Location/Qualifiers																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																										
--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--

Db 1741 AGAAGGAACCAATGTGTTGGGAGCCTTGAGAGAGAGCATGGATAGGTTCTCACTGCGC 1800

QY 1792 TCACAGGAGCAAGAAAGGGTTCTGCCAGAGACGGGCCGACAGAGCAATAAGCTTCACT 1851

Db 1801 TCACAAGGAACAAGAAAGGGCTCTGCCAGAGATGGACCAAGAAAGCGAAAGCTGCACT 1860

QY 1852 ATAAATGTGACTACAACTAGATTAAGTGAATTCAGATCACTGTTGTAATGAATTAATGTCTA 1911

Db 1861 ACAATGTGACTACAACTGCGCTGGTGAACACACACAGCTCTCTGAGCGGAATCATGGCTA 1920

QY 1912 TTCTTCCAAAGAGCAGTGTGACTTTGTACAAAAGGTTATACACTGAAGTGTCAAAACAC 1971

Db 1921 TTCTTCCAAAGAGACGTGGACTCTGTACAGAAAGTTACACTTAAGTGTCAAAAGC 1980

QY 1972 AGTCAGATTTGGAAAAGTGACAAATGCAATTTGAATTTAGAAGTGTCCAGCTTCAAAAC 2031

Db 1981 AGTCTGATTTGGCAAAGTGACAATGCAGTTTGAACGTGGAAGTGTGCCAGCTGCAGAGAC 2040

QY 2032 CCGATGTGGTGGTATCAGAGGACGCGCTTAAGGGCGATGCTGGTGTTCACAAAGAT 2091

Db 2041 CTGACGTGGTAGGCATCCGAGACAGCGCTGAAGGGTGTATGCTGGTGTTCACAGAGAT 2100

QY 2092 TAGTGAAGACATCTCTATCTAGCTGCAAGGTATATTTGATGGATTCTTCCATCTGCGCG 2151

Db 2101 TAGTGAAGATATCTTGTCTGCTGCAAGATGTGACCGATGACAGTGTCCGTGCGCGCG 2160

QY 2152 ATGAGTGTGGGTGTGATACAGCCTACATAAAGACTGTATATGATCGCTTTGATT-----T 2205

Db 2161 ATGAATGTGGTGTGATTTGACACAGGAGACTGTGAGATTGGTTCCATTTGGGGGCC 2220

QY 2206 TAAAGTTCATTTGAATCACTCACTTGTCTTAAAGAGTATCTTAAAGACCAATATCTCTT 2265

Db 2221 ATTGGATCTATGGAATCAAGCTCATT--TAAATAGCCATCTTCAAGACCAATGATTT 2278

QY 2266 TGTTTTAAACAAAGATATTTTGTGTATGAATCTAAATCAAGCCCATCTGTCATTA 2325

Db 2279 TTGCTTTAAACAAATTTTACTTTGTGGCTGAATCCAAGGCAA--CTGTTCTGTCATTA 2337

QY 2326 T--GTTACTGTCTTTTAAATCATGTGTTTTGTATATTAATAATTTGTGACTTTCTAG 2383

Db 2338 TCAATTACTCTGTTTAAATCATG--GGCTTGTATGTAGTAACGT-----TTTTAT 2390

QY 2384 ATTCACTTCCATATGTGAATGAAGCTTCAATATGCTCTTTGTAATGTGAATTTCT 2443

Db 2391 ATTCACTTCCATCAGGGAGTGCAGCTTCACTGTGACTCATTTTGTATGTACAGTTTCT 2450

QY 2444 TTCTGAAATAAAACCAATTTGTGAATAT 2470

Db 2451 TTCTGAACTAAACCAATTTGTGAATAT 2477

RESULT 8

BC020416

LOCUS BC020416.1 GI:18042795

DEFINITION Mus musculus maternal embryonic leucine zipper kinase, mRNA (cdna clone IMAGE:3588164).

ACCESSION BC020416

VERSION

KEYWORDS

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 2396)

AUTHORS Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, P., Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.P., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Udwin, T.B., Toshiyuki, S., Carninci, P., Frange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,

Abramson, R.D., Mullaby, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S.C., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalon, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahy, J., Helton, E., Kettman, M., Madan, A., Rodriguez, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smalls, D.E., Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.

Mammalian Gene Collection Program Team

Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences

Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

2 (bases 1 to 2396)

NIH MGC Project

Direct Submission

Submitted (03-JAN-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Bethesda, MD 20892-2590, USA

NIH-MGC Project URL: <http://mgc.nci.nih.gov>

Contact: MGC help desk

Email: [cgapbs@mail.nih.gov](mailto:cgapbs@mail.nih.gov)

Tissue Procurement: Gilbert Smith, Ph.D.

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Baylor College of Medicine Human Genome Sequencing Center

Center code: BCM-HGSC

Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>

Contact: [amg@bcm.tmc.edu](mailto:amg@bcm.tmc.edu)

Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Louesged, H., Kowis, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Nanavati, A.N., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>

Series: IRAK plate: 10 Row: b Column: 2

This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 6680605

This clone has the following problem: no 5' EST match.

Location/Qualifiers

1. 2396

/organism="Mus musculus"

/mol\_type="mRNA"

/strain="Czech II"

/db\_xref="taxon:10090"

/clone="IMAGE:3588164"

/tissue type="Mammary tumor metastasized to lung. Tumor arose spontaneously from a senescent normal mammary (clonal) outgrowth infected with the virus MMTV."

/clone\_lib="NCI CGAP\_Lu29"

/lab\_host="DH10B"

/note="Vector: pCMV-SPORT6"

Query Match 61.8%; Score 1525.8; DB 6; Length 2396;

Best Local Similarity 80.6%; Pred. No. 0;

Matches 1884; Conservative 0; Mismatches 412; Indels 41; Gaps 7;

QY 143 TAAATCCAAATAAACTTCCAGAGAGACTATGAAGATTATGATGAACCTTCTCAAAATATTA 202

Db 2 TCAATTTGATTAACCTTCCAGGTGGACTATGAAGATTATGACGAACCTCTCAATACTA 61

QY 203 TGAATTAATGAATATTTGGGACAGAGTGGCTTTTCCAAAGGTCAAACTTGCCTGCATAT 262

Db 62 TGAATATATGAAGAGTGGGACAGAGTGGCTTTTCCAAAGGTCAAACTTGCCTGCATAT 121

QY 263 CCTTACTGAGAGATGGTAGCTATATAAAATCATGATATAAAACACACACTAGGAGGTATT 322

Db 122 CCTCACTGAGAGATGGTAGCTATATAAAATCATGATATAAAATGCGCTAGGAGGTATT 181



AK137000	AK137000	2559 bp	mrna	linear	HTC 21-SEP-2005
LOCUS	Mus musculus 12 days embryo embryonic body between diaphragm region and neck cDNA, RIKEN full-length enriched library, clone:9430031C20 product:maternal embryonic leucine zipper kinase, full insert sequence.				
ACCESSION	AK137000	1	GI:74205778		
VERSION	AK137000.1				
KEYWORDS	HTC; CAP trapper.				
SOURCE	Mus musculus (house mouse)				
ORGANISM	Mus musculus				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muroidae; Muridae; Murinae; Mus.				
AUTHORS	1				
TITLE	Carninci, P. and Hayashizaki, Y.				
JOURNAL	High-efficiency full-length cDNA cloning				
PUBLISHED	Meth. Enzymol. 303, 19-44 (1999)				
REFERENCE	10349636				
AUTHORS	2				
TITLE	Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.				
JOURNAL	Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes				
PUBLISHED	Genome Res. 10 (10), 1617-1630 (2000)				
REFERENCE	11042159				
AUTHORS	3				
TITLE	Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsuina, T., Tashiro, H., Itoh, M., Sumi, N., Iehi, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.				
TITLE	RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer				
JOURNAL	Genome Res. 10 (11), 1757-1771 (2000)				
PUBLISHED	11076861				
REFERENCE	4				
AUTHORS	Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y., Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S., Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamanaka, I., Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R., Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T., Fletschmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H., Kuehl, P., Lewis, S., Matsuo, Y., Nikaido, I., Pesole, G., Quackenbush, J., Schriml, L.M., Staubli, P., Suzuki, R., Tomita, M., Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Bono, H., Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N., Carninci, P., de Bona, M.F., Brownstein, M.J., Bult, C., Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D., Hofmann, M., Hume, D.A., Kamiya, M., Lee, N.H., Lyons, P., Marchionni, L., Mashima, J., Mazzarelli, J., Mombaerts, P., Nordone, P., Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H., Sato, K., Schonbach, C., Seta, F., Shibata, Y., Storch, K.F., Suzuki, H., Toyooka, K., Wang, K.H., Weitz, C., Whitaker, C., Wilming, L., Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawai, H., Kohtsuki, S. and Hayashizaki, Y.				
CONSRMT	RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium				
TITLE	Functional annotation of a full-length mouse cDNA collection				
JOURNAL	Nature 409 (6821), 685-690 (2001)				
PUBLISHED	11217851				
REFERENCE	5				
AUTHORS	Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S., Nikaido, I., Osato, N., Saito, R., Suzuki, H., Yamanaka, I., Kiyosawa, H., Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A., Schonbach, C., Gojobori, T., Baldarelli, R., Hill, D.P., Bult, C., Hume, D.A., Quackenbush, J., Schriml, L.M., Kanapin, A., Matsuda, H., Batalov, S., Beisel, K.W., Blake, J.A., Bradt, D., Brusic, V., Chochia, C., Corbani, L.E., Cousins, S., Dalla, E., Dragani, T.A., Fletcher, C.F., Forrest, A., Frazer, K.S., Gaasterland, T., Gariboldi, M., Gissi, C., Godzik, A., Gough, J., Grimmond, S.,				
CONSRMT	Katayama, S., Tomaru, Y., Kasukawa, T., Waki, K., Nakanishi, M., Nakamura, M., Nishida, H., Yap, C.C., Suzuki, M., Kawai, J., Suzuki, H., Carninci, P., Hayashizaki, Y., Welle, C., Frith, M., Ravasi, T., Pang, K.C., Hallinan, J., Mattick, J., Hume, D.A., Lipovich, L., Batalov, S., Engstrom, P.G., Mizuno, Y., Faghihi, M.A., Sandelin, A., Chalk, A.M., Mottagui-Tabar, S., Liang, Z., Lennhard, B. and Wahlestedt, C.				
TITLE	The transcriptome landscape of the mammalian genome				
JOURNAL	Science 309 (5740), 1559-1563 (2005)				
PUBLISHED	16141072				
REFERENCE	7				

CONSRMT  
TITLE  
JOURNAL  
PUBLISHED  
REFERENCE  
AUTHORS

CONSRMT  
TITLE  
JOURNAL  
PUBLISHED  
REFERENCE  
AUTHORS





Encyclopedia Project of Genome Exploration Research Group in Riken  
Genomic Sciences Center and Genome Science Laboratory in RIKEN.  
Division of Experimental Animal Research in Riken contributed to  
prepare mouse tissues.  
Please visit our web site for further details.  
URL: <http://genome.gsc.riken.jp/>  
URL: <http://fantom.gsc.riken.jp/>.

## FEATURES

source  
1. .2488  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="C57BL/6J"  
/db\_xref="FANTOM DB:2410015K02"  
/db\_xref="taxon:10090"  
/clone="2410015K02"  
/cell\_type="ES cells"  
/clone\_lib="RIKEN full-length enriched mouse cDNA library"  
misc\_feature  
1. .2488  
/notes="maternal embryonic leucine zipper kinase  
(MGL|MG1.106924)  
putative"

## ORIGIN

Query Match 61.6%; Score 1520.4; DB 6; Length 2488;  
Best Local Similarity 80.4%; Pred. No. 0;  
Matches 1905; Conservative 0; Mismatches 421; Indels 42; Gaps 9;  
QY 112 GCGTGCCTCCGCCCTCAGGTTCTTTTCTAAATTCGAAATTAACCTTGCAGAGACTA 171  
DB 103 GCTGGGCCCGCTCCCTTCAGGTTCTGGTTGTCTATTTTGAATTAACCTTCCAGGTGGACTA 162  
QY 172 TGAAGATTATGATGAACCTCTCAATATTTACATGAATTAACCTTGGGACAGGTG 231  
DB 163 TGAAGATTATGACGAACCTCTCAATATCTATGAACTATATGAACGATTTGGACAGGTG 222  
QY 232 GCTTTGCCAAGGTCACAACTTGCCTGCCATATCTTACTGGAGAGATGGTAGCTATAAAAA 291  
DB 223 GCTTTGCCAAGGTCACAACTTGCCTGCCATATCTTACTGGAGAGATGGTAGCTATAAAAA 282  
QY 292 TCATGATATAAACAACACTAGGAGTGAATTTGCCCGGATCAAAACGGAGATTTAGGCGCT 351  
DB 283 TCATGATATAAACAACACTAGGAGTGAATTTGCCCGGATCAAAACGGAGATTTAGGCGCT 342  
QY 352 TGAAGAACCTTGACATACAGCATATATGCACTTACCATGCTGCTAGAGACAGCCACA 411  
DB 343 TGAAGAGCTTGACATACAGCATATGTCAGCTCTACCATGCTGCTGAGACAGCAAGACA 402  
QY 412 AAATATTATGTTCTTTGAGTACTGCCCTGGAGGAGCTGTTTGTACTATATAATTTCCC 471  
DB 403 AAATATTATGTTCTTTGAGTACTGCCCTGGAGGAGCTGTTTGTACTATATAATTTCCC 462  
QY 472 AGGATCGCTGTGAGAGAGGACACCGGGTGTCTTCGTGATAGTATCTGCTGTTG 531  
DB 463 AGGATCGCTGTGAGAGAGGACACCGGGTGTCTTCGTGATAGTATCTGCTGTTG 522  
QY 532 CTATTGTGACAGCCAGGCTATGCTCACAGGACCTCAAGCCAGAAAAATTTGCTGTTG 591  
DB 523 CGTATGTCCACAGCCAGGCTATGCTCACAGGACCTCAAAACAGAAAAATTTATTTATTTG 581  
QY 592 ATGAATATCATAAATTAAGCTGATTGACTTTTGGTCTCTGTGCAAAACCCCAAGGGTAACA 651  
DB 582 ATGAATATCATAAATTAAGCTGATTGACTTTTGGTCTCTGTGCAAAACCCCAAGGGTAACA 641  
QY 652 AGGATACCATCTACAGACATGCTGTGGAGTCTGGCTTATGAGACACCTGAGTTAATAC 711  
DB 642 AGGATACCATCTACAGACATGCTGTGGAGCTTGGCTTATGAGACACCTGAGTTAATAC 701  
QY 712 AAGCAAAATCATATCTTGGATCAGAGCAGATCTTTGGAGCTGGGACTCTGTTATATG 771  
DB 702 AAGGGAAGTCGTACCTTGGATCAGAGCAGATCTTTGGAGCTGGGACTCTGTTATATG 761  
QY 772 TTCTTATGTGTGATTTCTACCAATTTGATGATGAATGTAAATGGCTTTTACAAGAAGA 831  
DB 772 TTCTTATGTGTGATTTCTACCAATTTGATGATGAATGTAAATGGCTTTTACAAGAAGA 831

DB 762 TGCTCATGTGTGGATTTCTACCAATTTGATGATGAATGTATGTTCATAGCAAGAAGA 821  
QY 832 TTATCAGAGGAAAAATATGATGTTCCCAAGTGCTCTCTCCCAAGTAGCATTTCTGCTTCTTC 891  
DB 822 TAATCAGAGGAAAAATACGAAGTTCTTAAGTGCTCTCTCCCAAGTAGCATTTCTGCTTCTTC 881  
QY 892 AACAAATGCTGCAGGTGGACCCCAAGAAACGGATTTCTATGAAAAATCTATTGAACCATC 951  
DB 882 AGCAGATGTTGCAGGTGGACCCCAAGAAACGGATTTCTATGAAAAATCTCTCTGAACCATC 941  
QY 952 CTGATCATGCAAGATTATCAATCTATCTGTTGAGTGGCAAGCAAGATCTCTTTTATTC 1011  
DB 942 CTTGGGTATGCAAGATTATCAGCTGTCCGTGGAGTGGCAAGCAAGACTCTCTTTGACTC 1001  
QY 1012 ACCTCGATGATGATTGCGTAACAGAACTTTCTGTATCATCACAGAAACAAACGCAACAA 1071  
DB 1002 ACCTCGATGAGATTGCGTGACAGACTTTCTGTATCATCCGACGACGACGACGACAA 1061  
QY 1072 TGGAGGATTTAATTTCACTGTGGCAGTATGATCACCCTCAACGCTACCTATCTTCTGCTTC 1131  
DB 1062 TGGAGGATTTAATTTCTGTCGTGGCAGTACGATCACTCACAGCCACCTACCTTCTGCTTC 1121  
QY 1132 TAGCCAAAGGCTCGGGGAAAAACAGTTCTGTTTAAGGCTTTCTTCTTCTCTGCTGGAC 1191  
DB 1122 TAGCCAAAGGCTCGGGGAAAAACGCTGCTTACAGCTCTCTTCTCTTCTTCTGTTGAA 1181  
QY 1192 AAGCCAGTCTACCCCAATTCACAGACATCAAGTCAAAATAATTTGGAGTCTGGAAGATGCA 1251  
DB 1182 CGCCAGACACCACTCCA-----AAGTCAAGAAATCTGAGCTTGGAGATATGA 1229  
QY 1252 CGCAAGTGATAAAAATTTATGTGGGGGATTAATAGACTATATGATTTGGTGTGAAGATGAT 1311  
DB 1230 GCACAGTGATGATTAATCTGTGTGGCTGGATTGTAGACTATGAAATTTGTGTGAAGATAA 1289  
QY 1312 TATCAACAGGTGCTGCTACTCCCGAACAATCAAGTTTACCAAGTACTGGACAGAAATCAA 1371  
DB 1290 TATTA-----GCTCCCAAGACGCCACAGGTACCAAAACACTTTGGCAGAAATCAA 1337  
QY 1372 ATGGGTTGAAATCAATCAATTAATCCAGCTTATGCAAGAACCTGCAAAATAAATTA 1431  
DB 1338 ATCAGCGAGACTCTAAATCACCAGCGCCAGGGTACGACAGACAGTGGCAAAATAAATTA 1397  
QY 1432 AGAACAAAGAAATGTATATCTCTTAAGTCTGCTGTAAAGAAATGAAGAGTACTTTATGT 1491  
DB 1398 TGAACAAAGAAATGTGTGCACTCCCAAGTCTTCTGTGAAGAAATGAAGACAGTTTGTAT 1457  
QY 1492 TTCCTGAGCAAGACTCCAGTTTATAGAACCCAGCATAGAGAGAAATACTCACTACGC 1551  
DB 1458 TTTCTGAGCCGAAGATTCCAGTTAGTAAAGAACCCAGTATAAGAGAGAAATACCCGCTCAC 1517  
QY 1552 CAAATCGTTTACACTACACCTCAAAAGCTAGAAAAACAGTGCCTGAAAGAAACTCCAAATTA 1611  
DB 1518 CAAACCGTTTCCAAACACTGCAAAAGCTTAGAGCCAGTGCCTGAGAGAAAGCCCGGTTA 1577  
QY 1612 AAATACCAAGTAAATTAACAGGAAACAGCAAGTAAATGAACAGGTGTCAATTAGCCCTGAGA 1671  
DB 1578 GAACACCAAGGAAATTCGACAGGACAGACACACTAACGACAGGTGTCTATTAGCCCCGAGA 1637  
QY 1672 GCGGTGCGCTCAGTGGAAATTTGATCTCAACCAAGCAATATGAGGAGACTCCCAAAA 1731  
DB 1638 GGAGGTGCGCTTCAATGGACGTGGATCTCAACCCAGGCAACATGGAGGATATCCCGAAAA 1697  
QY 1732 GAAAGGGACCAAGTGTGTTGGAGCTTTGAAAGGGGTTGGATAGGTTTATCAGTGTGC 1791  
DB 1698 AGAAGGAACCAATGTGTTTGGAGCTTTGAGAGAGGACTGGATTAAGTTCTCAGTGC 1757  
QY 1792 TCACCAAGGACAAAAAGAGGGTTTCTGCCAGAGACGGGCCCCAGAAAGACTTAAAGCTTCACT 1851  
DB 1758 TCACCAAGGAACAAAGAAAGGGCTCTGCCAGAGATGGACCAAGAAAGCAAGCTCACT 1817  
QY 1852 ATATGTGATCAACACTAGATTAGTGAATTCAGATCACTGTTGAATGAATAATGCTTA 1911  
DB 1818 ACAATGTGATCAACTCGCTGGTGGTGAACCCCGACCAAGCTCTCTGAGCGAAATCATGGCTA 1877



QY	1912	TTCTTCCAAAGAGCATGTTGACTTGTGACAAAGGTTATACACTGAAGTGTCAAACA	1970
Db	1878	TTCTTCCAAAGAGAGCATGTTGACTTGTGACAAAGGTTATACACTGAAGTGTCAAACG	1937
QY	1971	CAGTCAGATTTTGGGAAGTGACAATGCAATTTGAATTTAGAACTGTGCCAGCTTCAAAA	2030
Db	1938	CAGTCAGATTTTGGGAAGTGACAATGCAATTTGAATTTAGAACTGTGCCAGCTGCAGAGA	1997
QY	2031	CCCGATGTGTGGGTATCAGGAGCGCAGCGCTTTAAGGGCGATGCCCTGGGTTTACAAAAGA	2090
Db	1998	CCTGAGCTGTGTAGGCTCCGGAGACAGCGCTGAAGGGTGATGCCCTGGGTTTACAGAGA	2057
QY	2091	TTAGTGGAGACATCTCTATCTAGCTGCAAGGTATAATTGATGATTTCTCCATCTGCGG	2150
Db	2058	TTAGTGGAGATATCTTGTCTGGCTGCAAGATGTGACCGATGACAGTGTCCGCTGCGCG	2117
QY	2151	GATGAGCTGTGGGTGATACAGCTCATATAAGCACTGTTATGATCGCTTTGATT-----	2204
Db	2118	GATGAAGGAGGTGTGATGTTGCACAGAGAGCTGTGAGATGTTGCCATTTGGGGGC	2177
QY	2205	TTAAAGTCTCATTTGGAATCACTCACTTTGTTTCTAAGAGAGCTATCTTAAGACCAATATCTCT	2264
Db	2178	CATTGATCTATGGAATCACTCACTTTGTTTCTAAGAGAGCTATCTTAAGACCAATATTTT	2237
QY	2265	TTGTTTTTAAACAAAGATATATTTTGTGTATGAATCTAAATCAAGCCCATCTGTCAAT	2324
Db	2238	TGCTTTTTTAAACAAAGATATATTTTGTGTATGAATCTAAATCAAGCCCATCTGTCAAT	2296
QY	2325	AT--GTTACTGCTTTTAAATCATGCTGTTTGTATATTAATTTGTTGACTTTCCTTA	2382
Db	2297	ATCAATCTCTCTGTTTTTAAATCATG--GGCTGTGTATGTAGTAACTG-----TTTTTA	2349
QY	2383	GATTCACTTCCATATGTAATGTAAGCTCTTAATCATGCTCTTTGTAATGTATTAATTTTC	2442
Db	2350	TATTCACTTCCATCAGGGATGCCAGCTCTTCACTGTGACTCATTTTGTATGATACAGTTTC	2409
QY	2443	TTTCTGAAATAAAACCATTTGTGAATAT	2470
Db	2410	TTTCTGAACTAAACCATTTGTGAATAT	2437
RESULT 11			
DQ048096		1956 bp	DNA
LOCUS			linear
DEFINITION			GSS 02-JUN-2005
ACCESSION			Homo sapiens MELK gene, VIRTUAL TRANSCRIPT, partial sequence,
VERSION			genomic survey sequence.
KEYWORDS			DQ048096
SOURCE			DQ048096.1 GI:66901295
ORGANISM			GSS.
REFERENCE			Homo sapiens (human)
AUTHORS			Homo sapiens
TITLE			Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
JOURNAL			Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
PUBLISHED			Hominidae; Homo.
REFERENCE			1 (bases 1 to 1956)
AUTHORS			Nielsen, R., Bustamante, C., Clark, A.G., Glanowski, S., Sackton, T.B.,
TITLE			Hubisz, M.J., Fedel-Alon, A., Tanenbaum, D.M., Cividello, D.,
JOURNAL			White, T.J., Sninsky, J.J., Adams, M.D. and Cargill, M.
COMMENT			A Scan for Positively Selected Genes in the Genomes of Humans and
			Chimpanzees
			(er) PLoS Biol. 3 (6), E170 (2005)
			15869325
			2 (bases 1 to 1956)
			Nielsen, R., Bustamante, C., Clark, A.G., Glanowski, S., Sackton, T.B.,
			Hubisz, M.J., Fedel-Alon, A., Tanenbaum, D.M., Cividello, D.,
			White, T.J., Sninsky, J.J., Adams, M.D. and Cargill, M.
			Direct Submission
			Submitted (05-MAY-2005) Celera Genomics, 45 West Gude Drive,
			Rockville, MD 20850, USA
			This sequence was made by sequencing genomic exons and ordering
			them based on alignment. Translation starts at the beginning of
			alignment.

FEATURES		Location/Qualifiers	
source		1..1956	
gene		/organism="Homo sapiens"	
		/mol_type="genomic DNA"	
		/db_xref="taxon:9606"	
		/chromosome="9"	
		<1..>1956	
		/gene="MELK"	
		/locus_tag="HC16044"	
ORIGIN			
Query Match		55.4%;	Score 1368; DB 14; Length 1956;
Best Local Similarity		69.9%;	Pred. No. 0;
Matches 1368; Conservative		0;	Mismatches 588; Indels 0; Gaps 0;
QY	171	ATGAAGAATTATGATGAACCTTCTCAAAATATTATGAATTACATGAACAACTATTGGGACAGGT	230
Db	1	ATGAAGAATTATGATGAACCTTCTCAAAATATTATGAATTACATGAACAACTATTGGGACAGGT	60
QY	231	GGCTTTGCAAAAGTCAAACTTGCCTGCCATATCCTTACTGGAGAGATGTAGCTATAAAA	290
Db	61	GGCTTTGCAAAAGTCAAACTTGCCTGCCATATCCTTACTGGAGAGATGTAGCTATAAAA	120
QY	291	ATCATGGATATAAAACACACTAGGAGTGTATTTGCCCGGATCAAAACGAGATTTAGGGCC	350
Db	121	ATCATGGATATAAAACACACTAGGAGTGTATTTGCCCGGATCAAAACGAGATTTAGGGCC	180
QY	351	TTGAAGAACTCAGACATCAGCATATATGTCAACTCTACCATGTGCTAGACAGACGCAAC	410
Db	181	NN	240
QY	411	AAAAATATCATGGTTCTTGAGTACTGCCCTGGAGAGAGCTGTTTGACATATAAATTTCC	470
Db	241	NN	300
QY	471	CAGGATCGCTGTCTGAGAGAGAGACCCGGGTGTTCTTCGGTCAGATAGTATCTGCTGTT	530
Db	301	CAGGATCGCTGTCTGAGAGAGAGACCCGGGTGTTCTTCGGTCAGATAGTATCTGCTGTT	360
QY	531	GCTTATGTGCACAGCGGCTATGCTCACAGGACCTCAAGCCAGAAATTTGCTGTTT	590
Db	361	GCTTATGTGCACAGCGGCTATGCTCACAGGACCTCAAGCCAGAAATTTGCTGTTT	420
QY	591	GATGAATATCAATAAATTAAGCTGATGACTTTGGTCTCTGTGCAAAACCCCAAGGGTAAC	650
Db	421	GATGAATATCAATAAATTAAGCTGATGACTTTGGTCTCTGTGCAAAACCCCAAGGGNNNN	480
QY	651	AAGGATTACCATCTACAGACATGCTGTGGAGTCTGGCTTATGCAGACCTCAGTTAATA	710
Db	481	NN	540
QY	711	CAAGGCAAAATCATATCTTGGATCAGAGGAGATGTTTGGAGCATGCGGCATCTGTTTATAT	770
Db	541	NN	600
QY	771	GTCTTATGTGTGGATTTCTACCATTTGATGATATAATGTAATGGCTTTATACAAGAAG	830
Db	601	NN	660
QY	831	ATTATGAGAGGAAATATGATGTTCCCAAGTGGCTCTCTCCAGTAGCATTCTGCTCTT	890
Db	661	NN	720
QY	891	CAACAAATGCTGCGATGGAGCCCAAGAAACCGATTTCTATGAAATACTTATGTAACCAT	950
Db	721	NN	780
QY	951	CCCTGGATCATGCAAGATTACAACTATCTGTGAGTGGCGCAAGCAAGATCTTTTATT	1010
Db	781	CCCTGGATCATGCAAGATTACAACTATCTGTGAGTGGCGCAAGCAAGATCTTTNNNNN	840
QY	1011	CACCTCGATGATGTTGGTAAACAGAACTTTCTGTACATCACAGAAACCAAGGCAACA	1070



```
QY 938 TCTATTGAACATCCCTGGATCATGCAAGATTACAACTATCCTGTTGAGTGGCAAGCAA 997
Db 738 TCTATTGAACATCCCTGGATCATGCAAGATTACAACTATCCTGTTGAGTGGCAAGCAA 679
QY 998 GAATCCTTTTATTACCTCGATGATGATGCGTAACAGAACTTCTGTACATCACAGAA 1057
Db 678 GAATCCTTTTATTACCTCGATGATGATGCGTAACAGAACTTCTGTACATCACAGAA 619
QY 1058 CAACAGCAAAACAAATGAGGAGTTAAATTCACCTGTGGCAGTATGATCACCTCACGGCTAC 1117
Db 618 CAACAGCAAAACAAATGAGGAGTTAAATTCACCTGTGGCAGTATGATCACCTCACGGCTAC 559
QY 1118 CTATCTCTCTCTAGCCAAAGAGGCTCGGGGAAACACAGTTGCTGTTTAAAGCTTTCTTC 1177
Db 558 CTATCTCTCTCTAGCCAAAGAGGCTCGGGGAAACACAGTTGCTGTTTAAAGCTTTCTTC 499
QY 1178 TTTCTCTCTGGCAAGCCAGTGTACCCCATTCACAGACATCAAGTCAAAATATTGGAG 1237
Db 498 TTTCTCTCTGGCAAGCCAGTGTACCCCATTCACAGACATCAAGTCAAAATATTGGAG 439
QY 1238 TCTGGAAGATGTGACCGCAGTGTATATAAAATTTATGTGGCGGGATTAAATAGACTATGATTG 1297
Db 438 TCTGGAAGATGTGACCGCAGTGTATATAAAATTTATGTGGCGGGATTAAATAGACTATGATTG 379
QY 1298 GTGTGAAGATGATTTATCAACAGGTGTGCTACTCTCCCGAAACATCAAGTTTACCAAGTA 1357
Db 378 GTGTGAAGATGATTTATCAACAGGTGTGCTACTCTCCCGAAACATCAAGTTTACCAAGTA 319
QY 1358 CTGACAGATCAATGGGGTGGAACTTAATCAATCAATTAATCTCCAGCCTTATGCAAGACAC 1417
Db 318 CTGACAGATCAATGGGGTGGAACTTAATCAATCAATTAATCTCCAGCCTTATGCAAGACAC 259
QY 1418 TGCAAAATAAATAAGAAACAAAGAAATGTATATCTCTTAAGTCTGCTGTTAAAGAAATGA 1477
Db 258 TGCAAAATAAATAAGAAACAAAGAAATGTATATCTCTTAAGTCTGCTGTTAAAGAAATGA 199
QY 1478 AGAGTACTTATGTTTCTGAGCCAAAGACTCCAGTTAATAAGAACAGCAGATTAAGAGAGA 1537
Db 198 AGAGTACTTATGTTTCTGAGCCAAAGACTCCAGTTAATAAGAACAGCAGATTAAGAGAGA 139
QY 1538 AATACCTACCTACGCCAAATGGTTTACACTACACCTCAAAAGCTAGAAACAGGCTCTGAA 1597
Db 138 AATACCTACCTACGCCAAATGGTTTACACTACACCTCAAAAGCTAGAAACAGGCTCTGAA 79
QY 1598 AGAACTCCAAATTAATAATACAGTAAATTCACAGCAAGACAGCAAGTTAATGACAGGTGT 1657
Db 78 AGTACTCCAAATTAATAATACAGTAAATTCACAGCAAGACAGCAAGTTAATGACAGGTGT 19
QY 1658 CATTAGCCCTCAGAGCG 1675
Db 18 CATTAGCCCTCAGAGCG 1

RESULT 13
AL554470 1030 bp mRNA linear EST 30-MAR-2004
LOCUS AL554470 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
DEFINITION clone CS0D1084Y107 5-PRIME, mRNA sequence.
ACCESSION AL554470
VERSION AL554470.3 GI:45859224
SOURCE EST.
ORGANISM Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1 (bases 1 to 1030)
Li, W.-B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished (2001)
On Feb 15, 2001 this sequence version replaced gi:31276280.
Contact: Genoscope
```

Genoscope - Centre National de Sequencage  
2 rue Gaston Cremlieux, CP 5706 - 91057 EVRY cedex - FRANCE  
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr  
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime  
end enriched, double-strand cDNA was digested with Not I and cloned  
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library  
was normalized. Library was constructed by Life Technologies, a  
division of Invitrogen. This sequence belongs to sequence cluster  
5403.f

For more information about this cluster, see  
http://www.genoscope.cns.fr/cdna?s=CS0D1084AE04QP1&c=5403.f.

## FEATURES

source

1..1030  
Location/Qualifiers  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="CS0D1084Y107"  
/tissue\_type="PLACENTA COT 25-NORMALIZED"  
/note="1st strand cDNA was primed with a NotI-oligo(dT)  
primer. Five prime end enriched, double-strand cDNA was  
digested with Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

## ORIGIN

Query Match 41.3%; Score 1020.6; DB 1; Length 1030;  
Best Local Similarity 99.5%; Pred. No. 3.9e-253;  
Matches 1020; Conservative 3; Mismatches 2; Indels 0; Gaps 0;  
QY 61 GCGCGTCTCTCAGGACAGCAGCGCCCTGCTCTCTGTCGGCGCGCGCTCAGCGGTGCC 120  
Db 1 GCGCGTCTCTCAGGACAGCAGCGCCCTGCTCTCTGTCGGCGCGCGCTCAGCGGTGCC 60  
QY 121 TCGCGCCCTCAGGTCTCTTTTCTTAATCCAAATAAATTCGCAAGAGGACTATGAAAAGATT 180  
Db 61 TCGCGCCCTCAGGTCTCTTTTCTTAATCCAAATAAATTCGCAAGAGGACTATGAAAAGATT 120  
QY 181 ATGATGAACCTTCTCAAAATATTAATGAATTAATGAAATCATGAACTATTGGGACAGGTGGCTTTGCAA 240  
Db 121 ATGATGAACCTTCTCAAAATATTAATGAATTAATGAAATCATGAACTATTGGGACAGGTGGCTTTGCAA 180  
QY 241 AGGTCAAACTTGCCTGCCATATCCTTACTGGAGAGATGTTAGTATATAAAATCATGAGATA 300  
Db 181 AGGTCAAACTTGCCTGCCATATCCTTACTGGAGAGATGTTAGTATATAAAATCATGAGATA 240  
QY 301 AAAACACACTAGGAGTGAATTTGCCCGATCAAAACGAGAGTTGAGGCTTTGAAGAACC 360  
Db 241 AAAACACACTAGGAGTGAATTTGCCCGATCAAAACGAGAGTTGAGGCTTTGAAGAACC 300  
QY 361 TGAGACATCAGCATATATGTCAACTCTACCATGTGCTAGAGACAGCCAAATAATTCA 420  
Db 301 TGAGACATCAGCATATATGTCAACTCTACCATGTGCTAGAGACAGCCAAATAATTCA 360  
QY 421 TGGTCTTGAGTACTGCTCGCTGGAGAGAGCTGTTTGAATATATATTTCCAGGATCGCC 480  
Db 361 TGGTCTTGAGTACTGCTCGCTGGAGAGAGCTGTTTGAATATATATTTCCAGGATCGCC 420  
QY 481 TGTCAAGAGAGAGACCGGGTGTCTTCCTCCGTCAGATAGTATCTGCTGTTGCTTATGTCG 540  
Db 421 TGTCAAGAGAGAGACCGGGTGTCTTCCTCCGTCAGATAGTATCTGCTGTTGCTTATGTCG 480  
QY 541 ACAGCCAGGCTATGCTCAAGGACCTCAAGCCAGAGAAATTTGCTGTTGATGAATATC 600  
Db 481 ACAGCCAGGCTATGCTCAAGGACCTCAAGCCAGAGAAATTTGCTGTTGATGAATATC 540  
QY 601 ATAAATTAAGCTGATGATGCTTGGTCTCTGTGCAAAACCCCAAGGTAACAGGATTACC 660  
Db 541 ATAAATTAAGCTGATGATGCTTGGTCTCTGTGCAAAACCCCAAGGTAACAGGATTACC 600  
QY 661 ATCTACAGACATGCTGTTGGAGTCTGGCTTATGACGACCTGAGTTATACAGGCAAT 720  
Db 601 ATCTACAGACATGCTGTTGGAGTCTGGCTTATGACGACCTGAGTTATACAGGCAAT 660

QY 721 CATATCTTGGATCAGAGGAGATGTTTGGAGCATGGGCATACCTGTTATATGTTCTTATGT 780  
DB |||||  
QY 781 GTGGATTTCTACCATTTGATGATGATAATGTAATGCTTTATACAGAGATTTATGAGAG 840  
DB |||||  
QY 841 GAAATATATGATGTTCCCAAGTGGCTCTCTCCAGTAGCATTTCTGCTTCTTCAACAAATGC 900  
DB |||||  
QY 901 TGCAGGTGGACCCCAAGAAACGGATTTCTATGAAAAATCTATTGAACCATCCCTGGATCA 960  
DB |||||  
QY 961 TGCAGGTGGACCCCAAGAAACGGATTTCTATGAAAAATCTATTGAACCATCCCTGGATCA 1020  
DB |||||  
QY 1021 ATGATTGCGTAAACAGAACTTCTGTATCATCACAGAAACCAACAGGCAAAACAATGGAGATT 1080  
DB |||||  
QY 1081 TAATT 1085  
DB |||||  
QY 1021 TATTT 1025

RESULT 14  
BM928784  
LOCUS  
DEFINITION  
AGENCOURT 6727232 NIH\_MGC\_100 Homo sapiens cDNA clone IMAGE:5798927  
5', mRNA sequence.  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
Homo sapiens (human)  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominidae; Homo.  
1 (bases 1 to 1091)  
NIH-MGC <http://mgi.nci.nih.gov/>.  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: [cgaps-k@mail.nih.gov](mailto:cgaps-k@mail.nih.gov)  
Tissue Procurement: CGAP (Stanford)  
cDNA Library Preparation: Rubin Laboratory  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
plate: LLCM2027 row: k column: 24  
High quality sequence stop: 708.

FEATURES  
source  
1. .1091  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5798927"  
/tissue\_type="hepatocellular carcinoma, cell line"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_lib="NIH\_MGC\_100"  
/notes="Organ: liver; Vector: pOTB7; Site:1; XhoI; Site 2:  
EcoRI; cDNA made by oligo-dT priming. Directionally cloned  
into EcoRI/XhoI sites using the following 5' adaptor:  
GGCACAG(G). Size-selected >500bp for average insert size  
1.8kb. Library constructed by Ling Hong in the laboratory  
of Gerald M. Rubin (University of California, Berkeley)  
using ZAP-cDNA synthesis kit (Stratagene) and Superscript

II RT (Life Technologies). Note: this is a NIH\_MGC  
Library."  
Query Match 35.3%; Score 872.4; DB 3; Length 1091;  
Best Local Similarity 98.3%; Pred. No. 1.1e-214;  
Matches 913; Conservative 0; Mismatches 12; Indels 4; Gaps 3;  
ORIGIN  
QY 899 GCTGCAAGTGGAGCCCAAGAAACGGATTTCTATGAAAAATCTATTGAACCATCCCTGGAT 958  
DB |||||  
QY 959 CATGCAAGATTACAACTATCTCTGTTAGTGGCAAGCAAGAAATCTTTTATTTCACCTCGA 1018  
DB |||||  
QY 1019 TGATGATTCGCTAAACAGACATTTCTGTACATCACAGAAACACAGCAACCAATGGAGGA 1078  
DB |||||  
QY 139 TGATGATTCGCTAAACAGAACTTTCTGTACATCACAGAAACACAGCAACCAATGGAGGA 198  
QY 1079 TTTAAATTTTCACTGTGGCAGTATGATCACCTCACGGCTACCTATCTTCTGCTTCTAGCCAA 1138  
DB |||||  
QY 199 TTTAAATTTTCACTGTGGCAGTATGATCACCTCACGGCTACCTATCTTCTGCTTCTAGCCAA 258  
QY 1139 GAAGGCTCGGGGAAACCAAGTTCTTTAAGGCTTTCTTCTCTCTGTGGCAAGCCAG 1198  
DB |||||  
QY 1199 TGCTACCCCATTTACAGACATCAAGTCAATTAATTTGGAGCTGGAGAGTGTGACCCGAG 1258  
DB |||||  
QY 319 TGCTACCCCATTTACAGACATCAAGTCAATTAATTTGGAGCTGGAGAGTGTGACCCGAG 378  
QY 1259 TGATAAAAATTTATGTGGCGGATTAATAGATCTATGATGGTGTGAAGATGATTTATCAAC 1318  
DB |||||  
QY 379 TGATAAAAATTTATGTGGCGGATTAATAGATCTATGATGGTGTGAAGATGATTTATCAAC 438  
QY 1319 AGGTGCTGCTACTCTCCCGAACATCAAGTTTACCAAGTACTGGACAGAAATCAATGGGGT 1378  
DB |||||  
QY 439 AGGTGCTGCTACTCTCCCGAACATCAAGTTTACCAAGTACTGGACAGAAATCAATGGGGT 498  
QY 1379 GGAATCTAAATCAATTAATCTCCAGCTTATGACAGACACCTGCANATATAATTAAGAACAA 1438  
DB |||||  
QY 499 GGAATCTAAATCAATTAATCTCCAGCTTATGACAGAACCTGCANATATAATTAAGAACAA 558  
QY 1439 AGAAATGTATATATCTCTTAAGTCTGCTTAAGATGAAGAGTACTTTATGTTTCTCTGA 1498  
DB |||||  
QY 559 AGAAATGTATATCTCTTAAGTCTGCTTAAGATGAAGAGTACTTTATGTTTCTCTGA 618  
QY 1499 GCCAAGACTCCAGTTTAAGAACCCAGCATTAAGAGAGAAATCTCACTACGCCAAATCG 1558  
DB |||||  
QY 619 GCCAAGACTCCAGTTTAAGAACCCAGCATTAAGAGAGAAATCTCACTACGCCAAATCG 678  
QY 1559 TTACACTACACCTCAAAAGCTAGAACAGCTGCCTGAAAGAACTCCCAATTAATAATACC 1618  
DB |||||  
QY 679 TTACACTACACCTCAAAAGCTAGAACAGCTGCCTGAAAGAACTCCCAATTAATAATACC 738  
QY 1619 AGTAAATTTCAACAGGAACAGACAAGTTAATGACAGGTGTCTTAGCCCTCAGAGCGGGTG 1678  
DB |||||  
QY 739 AGTAAATTTCAACAGGAACAGACAAGTTAATGACAGGTGTCTTAGCCCTCAGAGCGGGTG 798  
QY 1679 CCGCTCAGTGGAAATGGATCTCAACCAAGCATATGAGAGAGATCTCCAAAAGAAAGGG 1738  
DB |||||  
QY 799 CCGCTCAGTGGAAATGGATCTCAACCAAGCATATGAGAGAGATCTCCAAAAGAAAGGG 858  
QY 1739 AGCCAAAGTGTGTTGGAGCTTGAAG - -GGGGGTTGGAATGAAGTTATCACTGT -GCTCAC 1795  
DB |||||  
QY 859 AGCCAAAGTGTGTTGGAGCTTGAAG - -GGGGGTTGGAATGAAGTTATCACTGT -GCTCAC 1795  
DB |||||  
QY 1796 CAGGAGC -AAAAAGGAAGGGTTCCTGCCAGA 1823  
DB |||||  
QY 919 CAGGAGC -AAAAAGGAAGGGTTCCTGCCAAA 947  
DB |||||

```

RESULT 15
DQ048097
LOCUS
DEFINITION
  DQ048097 1898 bp DNA linear GSS 02-JUN-2005
  Pan troglodytes MELK gene, VIRTUAL TRANSCRIPT, partial sequence,
  genomic survey sequence.
ACCESSION
  DQ048097
VERSION
  DQ048097.1 GI:66901296
KEYWORDS
  GSS.
SOURCE
  Pan troglodytes (chimpanzee)
ORGANISM
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
  Homiidae; Pan.
REFERENCE
  1. (bases 1 to 1898)
  Nielsen, R., Bustamante, C., Clark, A.G., Glanowski, S., Sackton, T.B.,
  Hubisz, M.J., Fiedel-Alon, A., Tanenbaum, D.M., Civallo, D.,
  White, T.J., Sninsky, J.J., Adams, M.D. and Cargill, M.
  A Scan for Positively Selected Genes in the Genomes of Humans and
  Chimpanzees
  (er) PLoS Biol. 3 (6), E170 (2005)
  15869325
REFERENCE
  2. (bases 1 to 1898)
  Nielsen, R., Bustamante, C., Clark, A.G., Glanowski, S., Sackton, T.B.,
  Hubisz, M.J., Fiedel-Alon, A., Tanenbaum, D.M., Civallo, D.,
  White, T.J., Sninsky, J.J., Adams, M.D. and Cargill, M.
  Direct Submission
  Submitted (05-MAY-2005) Celera Genomics, 45 West Gude Drive,
  Rockville, MD 20850, USA
  This sequence was made by sequencing genomic exons and ordering
  them based on alignment. Translation starts at the beginning of
  alignment.
FEATURES
  source
    1. 1898
    /organism="Pan troglodytes"
    /mol_type="genomic DNA"
    /db_xref="taxon:9598"
  gene
    <1..>1898
    /gene="MELK"
    /locus_tag="HC16044"
  ORIGIN
    Query Match 34.8%; Score 859; DB 14; Length 1898;
    Best Local Similarity 45.4%; Pred. No. 4e-211;
    Matches 862; Conservative 0; Mismatches 1036; Indels 0; Gaps 0;

    229 GTGGCTTTGCAAGGTCACAACTTGCCTGCCATATCCTTACTGAGAGAGATGGTAGCTATAA 288
    |||||
    1 GTGGCTTTGCAAGGTCACAACTTGCCTGCCATATCCTTACTGAGAGAGATGGTAGCTATAA 60
    |||||

    289 AAATCATGATGATAAAACACACTAGGGAGTGATTTGGCCCGGATCAAAACGGAGATTGAGG 348
    |||||
    61 AAATCATGATGATAAAACACACTAGGGAGTGATTTGGCCCGGATCAAAACGGAGATTGAGG 120
    |||||

    349 CCTTGAAGAACCTGAGACATCAGCATATATGTCACTTACCATGTGCTAGACAGCCCA 408
    |||||

    121 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 180
    |||||

    409 ACAAAATATTTCATGGTCTTTCAGTACTGCTGAGGAGAGAGCTGTTGACTATATAATTT 468
    |||||
    181 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 240
    |||||

    469 CCAGGATCGCTGTGACAGAGAGAGACCCGGGTGTTCTTCCGTGAGATAGTATCTGCTG 528
    |||||
    241 CCAGGATCGCTGTGACAGAGAGAGACCCGGGTGTTCTTCCGTGAGATAGTATCTGCTG 300
    |||||

    529 TTGCTTATGTCAGCAGCCGGGTATGCTCAGAGGACCTCAAGCCAGAAATTTGCTGT 588
    |||||
    301 TTGCTTATGTCAGCAGCCGGGTATGCTCAGAGGACCTCAAGCCAGAAATTTGCTGT 360
    |||||

    589 TTGATGAATCATATAAATTAAGCTGATTGCTTGTCTCTGTGCAAAACCCCAAGGGTA 648
    |||||
    361 TTGATGAATCATATAAATTAAGCTGATTGCTTGTGCAAAACCCCAAGGGTTATCACTG 1788
    |||||
  
```

1501	AAAGAAAGGGAGCCAAAAGTGTGTTGGGAGCCCTGAAAGGGGTTGGATAAGGTTATCACATG	1560
Qy	TGCTCACGAGGAGCAAAAGGAGGGTCTGCCAGAGACGGGCCCCAGAGACTAAAGCTTC	1848
Db	TGCTCACGAGGAGCAAAAGGAGGGTCTGCCAGAGATGGGCCNNNNNNNNNNNNCTTC	1620
Qy	ACTATAATGTGACTACAACATAGATTAGTGAATCCAGATCAACTGTGTGAATGAAATAATGT	1908
Db	ACTATAATGTGACTACAACATAGATTAGTGAATCCAGATCAACTGTGTGAATGAAATAATGT	1680
Qy	CTATTCTTCAAAGAAAGCATGTTGACATTGTGTACAAAAGGGTTATACACTGAAAGTGTCAAA	1968
Db	CTATTCTTCAAAGAAAGCATGTTGACATTGTGTACNNAAGGGTTATACACTGAAAGTGTCAAA	1740
Qy	CACAGTCAGATTTTGGGAAAGTGACAATGCAATTTGAAATTTAGAAGTGTGCCAGCTTCAAA	2028
Db	CACAGTCAGATTTTGGGAAAGTGACAATGCAATTTTGAATTTAGAAGTGTGCCAGCTTCAAA	1800
Qy	AACCCGATGTGGTGGGTATCAGGAGGAGCGGGTTAAGGGCGATGCTCGTGGTTTACA AAA	2088
Db	AACCCGATGTGGTGGGTATCAGGAGGAGCGGGTTAAGGGCGATGCTCGTGGTTTACA AAA	1860
Qy	GATTAGTGGAAAGACATCCTATCTAGCTGCAAGGTATAA	2126
Db	GATTAGTGGAAAGACATCCTAGCTGCAAGGTATAA	1898

Search completed: November 21, 2006, 22:55:22  
Job time : 11529 secs

GenCore version 5.1.9  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 21, 2006, 19:41:18 ; Search time 2640 Seconds  
(without alignments)  
11496.386 Million cell updates/sec

Title: US-10-656-598-1

Perfect score: 2470

Sequence: 1 ttggcggggggaaggccca.....ataaacatttggatgaat 2470

Scoring table: IDENTITY NUC

Gapop 10\_0 , Gapext 1.0

Searched: 18922170 seqs, 6143817638 residues

Total number of hits satisfying chosen parameters: 37784340

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA Main:\*

- 1: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US07\_PUBCOMB.seq:\*
- 2: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US08\_PUBCOMB.seq:\*
- 3: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US09A\_PUBCOMB.seq:\*
- 4: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US09B\_PUBCOMB.seq:\*
- 5: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US09C\_PUBCOMB.seq:\*
- 6: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US10A\_PUBCOMB.seq:\*
- 7: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US10B\_PUBCOMB.seq:\*
- 8: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US10C\_PUBCOMB.seq:\*
- 9: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US10D\_PUBCOMB.seq:\*
- 10: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US10E\_PUBCOMB.seq:\*
- 11: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US10F\_PUBCOMB.seq:\*
- 12: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US10G\_PUBCOMB.seq:\*
- 13: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US11A\_PUBCOMB.seq:\*
- 14: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US11B\_PUBCOMB.seq:\*
- 15: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US11C\_PUBCOMB.seq:\*
- 16: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US11D\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2470	100.0	2470	3	US-09-870-937-9
2	2470	100.0	2470	3	US-09-954-456-1121
3	2470	100.0	2470	3	US-09-880-107-1725
4	2470	100.0	2470	7	US-10-354-358-7
5	2470	100.0	2470	7	US-10-172-118-1535
6	2470	100.0	2470	7	US-10-295-027-505
7	2470	100.0	2470	7	US-10-173-999-126
8	2470	100.0	2470	8	US-10-342-887-1535
9	2470	100.0	2470	9	US-10-656-598-1
10	2470	100.0	2470	9	US-10-723-860-2443
11	2470	100.0	2470	10	US-10-848-755A-169
12	2470	100.0	2470	10	US-10-843-641A-4148
13	2470	100.0	2470	10	US-10-756-149-2259
14	2470	100.0	2470	10	US-10-770-726-23
15	2470	100.0	2470	12	US-10-960-414-132
16	2470	100.0	2470	15	US-11-177-138-9
17	2470	100.0	2487	8	US-10-425-114-26285

18	2470	100.0	2360	9	US-10-723-860-6621	Sequence 6621, Ap
19	2452	99.3	2463	8	US-10-425-114-26959	Sequence 26959, A
20	2441	98.8	2463	8	US-10-425-114-26835	Sequence 26835, A
21	2438	98.7	2453	3	US-09-974-298-111	Sequence 111, App
22	2438	98.7	2501	10	US-10-955-054A-138	Sequence 138, App
23	2159.8	87.4	2519	6	US-10-071-766-82	Sequence 82, Appl
24	2127.6	86.1	2162	3	US-09-822-849A-483	Sequence 483, App
25	1398.4	56.6	1400	13	US-11-060-756-1002	Sequence 1002, Ap
26	1398.4	56.6	1400	13	US-11-060-756-5274	Sequence 5274, Ap
27	855.4	34.6	877	3	US-09-764-875-308	Sequence 308, App
28	843.2	34.1	887	6	US-10-106-698-1036	Sequence 1036, Ap
29	720.8	29.2	1673	3	US-09-814-353-21017	Sequence 21017, A
30	524	21.2	15965	15	US-11-121-086-56	Sequence 56, Appl
31	512	20.7	512	6	US-10-656-598-13	Sequence 13, Appl
32	507.6	20.6	606	6	US-10-106-698-2000	Sequence 2000, Ap
33	437.8	17.7	489	3	US-09-764-868-222	Sequence 222, App
34	437.8	17.7	489	3	US-09-764-875-585	Sequence 585, App
35	410	16.6	410	9	US-10-696-639-1095	Sequence 1095, Ap
36	385.8	15.6	423	3	US-09-918-995-33354	Sequence 33354, A
37	373	15.1	373	3	US-09-822-830A-155	Sequence 155, App
38	268.8	10.9	499	3	US-09-783-590-8126	Sequence 8126, Ap
39	256	10.4	496	3	US-09-783-590-7891	Sequence 7891, Ap
40	228	9.2	572	3	US-09-864-761-13539	Sequence 13539, A
41	228	9.2	572	13	US-11-060-756-1430	Sequence 1430, Ap
42	228	9.2	572	13	US-11-060-756-1431	Sequence 1431, Ap
43	228	9.2	572	13	US-11-060-756-5702	Sequence 5702, Ap
44	228	9.2	572	13	US-11-060-756-5703	Sequence 5703, Ap
45	208.4	8.4	1863	9	US-10-635-398-109	Sequence 109, App

ALIGNMENTS

RESULT 1  
US-09-870-937-9  
; Sequence 9, Application US/09870937  
; Patent No. US20020049180A1  
; GENERAL INFORMATION:  
; APPLICANT: Wu, Bin  
; APPLICANT: Seeley, Todd  
; APPLICANT: Williams, Lewis T.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING NEOPLASTIC DISEASE USING  
; FILE REFERENCE: 200130.514/PP-01623.002  
; CURRENT APPLICATION NUMBER: US/09/870,937  
; CURRENT FILING DATE: 2001-05-30  
; NUMBER OF SEQ ID NOS: 12  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 9  
; LENGTH: 2470  
; TYPE: DNA  
; ORGANISM: Homo sapien  
US-09-870-937-9

Query Match 100.0%; Score 2470; DB 3; Length 2470;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2470; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	TTGGCGGGCGGAGCGGCCACACCCGGCGATCGAAAGATTCTTAGAACCCGCTACCA	60
Db	1	TTGGCGGGCGGAGCGGCCACACCCGGCGATCGAAAGATTCTTAGAACCCGCTACCA	60
Qy	61	GC CGCGTCTCTCAGGACAGCAGGCCCTCTCTCTCTGTCGGCGCGCTCAGCGTGC	120
Db	61	GC CGCGTCTCTCAGGACAGCAGGCCCTCTCTCTCTGTCGGCGCGCTCAGCGTGC	120
Qy	121	TCCGCCCTCAGGTTCTTTTCTAATTCAAATAAACTTGAAGAGGACTTGAAGATT	180
Db	121	TCCGCCCTCAGGTTCTTTTCTAATTCAAATAAACTTGAAGAGGACTTGAAGATT	180
Qy	181	ATGATGAACCTCTCAATATATGATTAATGATTAATGATTAATGATTAATGATTAAT	240
Db	181	ATGATGAACCTCTCAATATATGATTAATGATTAATGATTAATGATTAATGATTAAT	240



QY	241	AGGTCAAACTTGGCTGCGCATATCCTTACTGAGAGATGGTAGCTATAAAAAATCATGGATA	300
Db	241		
QY	301	AAACACACTAGGAGGTGATTTGCCCGGATCAAAAACGGAGATTGAGGCTTCAAGAAC	360
Db	301		
QY	361	TGAGACATCAGCATATATGTCACCTCTACCATGTCTAGAGACAGCCACAAAATATTCA	420
Db	361		
QY	421	TGGTCTTTGAGTAGTCCCTCGGAGGAGCTGTTTGACTATATAATTTCCACGATCGCC	480
Db	421		
QY	481	TGTGAGAAGAGAGACCCGGTTGTCTTCGTCAGATAGTCTGCTGTTGCTTTATGTGC	540
Db	481		
QY	541	ACAGCCAGGCTATGCTCAAGGACCTCAAGCCAGAAAATTTGCTGTTGATGATATC	600
Db	541		
QY	601	ATAAATTAAGCTGATTGACTTTGGTCTCTGTGCAAAACCAAGGGTAACAAGGATTACC	660
Db	601		
QY	661	ATCTACAGACATGCTGAGGAGTCTGGCTTTATGACGACCTGAGTTAAATCAAGGCAAA	720
Db	661		
QY	721	CATATCTTGATCAGAGCAGATGTTTGGAGCATGGGCATACGTGTATATGTTCTTATGT	780
Db	721		
QY	781	GTGGATTTCTACCATTTGATGATGATACTGATGCTTTATACAAAGAGATTATGAGAG	840
Db	781		
QY	841	GAATAATATGATGTTCCCAAGTGGCTCTCCAGTAGCATCTGCTTTCTTCAACAAATGC	900
Db	841		
QY	901	TGCAAGTGGACCCAAAGAACGGATTCTATGAAAAATCTATTGAACCATCCCTGGATCA	960
Db	901		
QY	961	TGCAAGATTACAATCTATCTGTTGAGTGGCAAGCAAGAATCCTTTTATTCACCTCGATG	1020
Db	961		
QY	1021	ATGATTCGCTAACAGAACTTTCTGTATCATCACAGAAACAAAGCAAAATGGAGATT	1080
Db	1021		
QY	1081	TAATTTTCACTGTGGCAGTATGATCACCTCAGGCTACCTATCTTCTGCTTACGCCNAGA	1140
Db	1081		
QY	1141	AGGCTGGGGAACCAAGTTCTTTAAGGCTTTCTTTAAGGCTTTCTTCTCTGTGGCAAGCCAGTG	1200
Db	1141		
QY	1201	CTACCCCATTCACAGACATCAAGTCAAAATATGGAGTCTGGAAGATGTGACCGCAAGTG	1260
Db	1201		
QY	1261	ATAAATAATGTGGCGGATTAATAGACTATGATGATGTTGGTGTGAGATGATTTATCAACAG	1320
Db	1261		

QY	1321	GTGCTGCTACTCCCGAACATCACAGTTTAAACAAGTACTGGACAGAAATCAAAATGGGGTGG	1380
Db	1321		
QY	1381	AATCTAAATCATTAATCTCAGGCTTATGACAGAACACCTGCAAAATAAATTAAGAAACAAAG	1440
Db	1381		
QY	1441	AAAAATGATATACTCTTAAGTCTGCTGTAAAGAAATGAAGAGTACTTTATGTTTCCTGAGC	1500
Db	1441		
QY	1501	CAAAGACTCCAGTTATTAAGAACCCAGCATAAAGAGAGAAATCTCACTACGCCAAATCGTT	1560
Db	1501		
QY	1561	ACACTACACCTCAAAAGCTAGAAAACCAAGTCTGAAAGAAATCTCAATTAANAATCCAG	1620
Db	1561		
QY	1621	TAAATTTCAACAGCAAGCAAGTTAATGACAGGTGTCAATTAGCCCTGAGAGCGGTGCC	1680
Db	1621		
QY	1681	GCTCAGTGAATTTGGATCTCAACCAAGCACATATGGAGGAGACTCCAAAAGAAAGGGAG	1740
Db	1681		
QY	1741	CCAAAGTGTGGAGCCCTTGAAGGGGTTGGATAAGGTTATCACTGTGCTCACCAGGA	1800
Db	1741		
QY	1801	GCAAAAGGAAGGTTCTCGCAGAGACGGGCCACAGAGACTTAAAGCTTCACTATAATGTGA	1860
Db	1801		
QY	1861	CTACAACTAGATTAGTGAATCCAGATCAACTGTGTAATGAATAATGTCTATTCTTCCAA	1920
Db	1861		
QY	1921	AGAGCATGTTGACTTTGTACAAAAGGTTATACCTGAAGTGTCAACACAGTCAGATT	1980
Db	1921		
QY	1981	TTGGGAAAGTGACAAATTTGAATTAAGAGTGTCCAGCTTCAAAAACCCGATGTGG	2040
Db	1981		
QY	2041	TGGGTATCAGGAGGACAGCGCTTAAGGGCGATGCGTGGGTTTACAAAAGATTAGTGGAG	2100
Db	2041		
QY	2101	ACATCCTATCTAGCTGCAAGGTATATTGATGAGTTCTTCCATCCTGCCGATGAGTGTG	2160
Db	2101		
QY	2161	GGTGTATACAGCTCATATAAGAGCTGTTATGATCGCTTTGATTTTAAAGTTTCAATGGAA	2220
Db	2161		
QY	2221	CTACCAACTTGTGTTTAAAGAGCTATCTTAAGCAAAATATCTCTTTGTTTTTAAACAAA	2280
Db	2221		
QY	2281	GATATTTTGTGATGATGATCTAAATCAAGCCCATCTGTCTATTATGTTACTGCTTTTT	2340
Db	2281		
QY	2341	TAAATCATGTGTTTTGTATATTAATAATTTGACCTTTCTTAGATTCACTTCCATATGTG	2400
Db	2341		
QY	2401	AATGTAAGCTTTAACTATGTCTCTTTTGTAAATGTGTAATTTCTTTCTGAATAAACCAT	2460
Db	2401		

Db 2401 AATGTAAGCTCTTAACATATGCTCTCTTTGTAATGTAATTTCTTCTGAATAAAACCAT 2460  
QY 2461 TTGTGAATAT 2470  
Db 2461 TTGTGAATAT 2470  
RESULT 2  
US-09-954-456-1121  
; Sequence 1121, Application US/09954456  
; Patent No. US20020115057A1  
; GENERAL INFORMATION:  
; APPLICANT: Young, Paul  
; TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using Cand  
; FILE OF INVENTION: Sets  
; FILE REFERENCE: 689290-76  
; CURRENT APPLICATION NUMBER: US/09/954,456  
; PRIOR FILING DATE: 2001-09-18  
; PRIOR APPLICATION NUMBER: US/60/233,617  
; PRIOR FILING DATE: 2000-09-18  
; PRIOR APPLICATION NUMBER: US/60/234,052  
; PRIOR FILING DATE: 2000-09-20  
; PRIOR APPLICATION NUMBER: US/60/234,923  
; PRIOR FILING DATE: 2000-09-25  
; PRIOR APPLICATION NUMBER: US/60/235,134  
; PRIOR FILING DATE: 2000-09-25  
; PRIOR APPLICATION NUMBER: US/60/235,637  
; PRIOR FILING DATE: 2000-09-26  
; PRIOR APPLICATION NUMBER: US/60/235,638  
; PRIOR FILING DATE: 2000-09-26  
; PRIOR APPLICATION NUMBER: US/60/235,711  
; PRIOR FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: US/60/235,720  
; PRIOR FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: US/60/235,840  
; PRIOR FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: US/60/235,863  
; PRIOR FILING DATE: 2000-09-27  
; NUMBER OF SEQ ID NOS: 2276  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 1121  
; LENGTH: 2470  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-954-456-1121  
Query Match 100.0%; Score 2470; DB 3; Length 2470;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2470; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 TTGGCGGGCGGAAGCGGCCCAACCCGGCGATCGAAAGATTCTTAGGAACGCCGTACCA 60  
Db 1 TTGGCGGGCGGAAGCGGCCCAACCCGGCGATCGAAAGATTCTTAGGAACGCCGTACCA 60  
QY 61 GCGCGTCTCTCAGGACAGAGCCGCCCTCTCTTCTGTGCGGCGCGCTCAGCCGTCGCC 120  
Db 61 GCGCGTCTCTCAGGACAGAGCCGCCCTCTCTTCTGTGCGGCGCGCTCAGCCGTCGCC 120  
QY 121 TCGGCCCTCAGGTTCTTTTCTAATTCGAATAAATCTTGAAGAGGACTATGAAGATT 180  
Db 121 TCGGCCCTCAGGTTCTTTTCTAATTCGAATAAATCTTGAAGAGGACTATGAAGATT 180  
QY 181 ATGATGAACCTCTCAATATATTAATTAATGAACTATTCGGGACAGGTGCTTTGCCAA 240  
Db 181 ATGATGAACCTCTCAATATATTAATTAATGAACTATTCGGGACAGGTGCTTTGCCAA 240  
QY 241 AGGTCAAACTTCGCTGCCATATCTTCTACTGTGAGAGATGTTAGCTATAAAATCATGGATA 300  
Db 241 AGGTCAAACTTCGCTGCCATATCTTCTACTGTGAGAGATGTTAGCTATAAAATCATGGATA 300  
QY 301 AAAACACACTAGGAGTGTATTTGCCCGGATCAAAACCGAGATTGAGGCTTTGAAGAACC 360  
Db 301 AAAACACACTAGGAGTGTATTTGCCCGGATCAAAACCGAGATTGAGGCTTTGAAGAACC 360

301 AAAACACACTAGGAGTGTATTTGCCCGGATCAAAACCGAGATTGAGGCTTTGAAGAACC 360  
QY 361 TGAGACATCAGCATATATGTCAACTCTACATGTCTAGAGACAGCAAAATATTCA 420  
Db 361 TGAGACATCAGCATATATGTCAACTCTACATGTCTAGAGACAGCAAAATATTCA 420  
QY 421 TGGTTCTTGTAGTACTGCCCTGGAGGAGAGCTGTTTGAATTAATTTCCAGGATCGCC 480  
Db 421 TGGTTCTTGTAGTACTGCCCTGGAGGAGAGCTGTTTGAATTAATTTCCAGGATCGCC 480  
QY 481 TGTCAAGAGAGAGACCCGGGTTGTCTCCGTACAGATAGTATCTGCTGTCTTATGTGC 540  
Db 481 TGTCAAGAGAGAGACCCGGGTTGTCTCCGTACAGATAGTATCTGCTGTCTTATGTGC 540  
QY 541 ACAGCCAGGCTATGCTCAAGGGACCTCAAGCCAGAGAAATTTGCTTTGATGAATATC 600  
Db 541 ACAGCCAGGCTATGCTCAAGGGACCTCAAGCCAGAGAAATTTGCTTTGATGAATATC 600  
QY 601 ATAAATTAAGCTGATTGACTTTGGTCTCTGTGCAAAACCCAGGGTAACAGGATTACC 660  
Db 601 ATAAATTAAGCTGATTGACTTTGGTCTCTGTGCAAAACCCAGGGTAACAGGATTACC 660  
QY 661 ATCTACAGACATGCTGTGGAGTCTGGCTTATGACAGCACCTGAGTTAATACAAAGCAAT 720  
Db 661 ATCTACAGACATGCTGTGGAGTCTGGCTTATGACAGCACCTGAGTTAATACAAAGCAAT 720  
QY 721 CATATCTTTGGATCAGAGGCAGATGTTGGAGCATGGGCATCTGTTATATGTTCTTATGT 780  
Db 721 CATATCTTTGGATCAGAGGCAGATGTTGGAGCATGGGCATCTGTTATATGTTCTTATGT 780  
QY 781 GTGGATTTCTACCATTTGATGATGATTAATGTAATGGCTTTTATACAAAGAGATTATGAGAG 840  
Db 781 GTGGATTTCTACCATTTGATGATGATTAATGTAATGGCTTTTATACAAAGAGATTATGAGAG 840  
QY 841 GAAATATGATGTTCCCAAGTGGCTCTCCAGTAGCATTTCTGCTTCTTCAACAAATGTC 900  
Db 841 GAAATATGATGTTCCCAAGTGGCTCTCTCCAGTAGCATTTCTGCTTCTTCAACAAATGTC 900  
QY 901 TGCAGGTGGACCCAAAGAAACGGATTCTTATGAAATCTATTGAACCATCCCTGGATCA 960  
Db 901 TGCAGGTGGACCCAAAGAAACGGATTCTTATGAAATCTATTGAACCATCCCTGGATCA 960  
QY 961 TGCAGGATTACAATCTATCTGTGTGAGTGGCAAGCAAGATCCTTTTATACCTCGATG 1020  
Db 961 TGCAGGATTACAATCTATCTGTGTGAGTGGCAAGCAAGATCCTTTTATACCTCGATG 1020  
QY 1021 ATGATTGCGTAACGAACTTTCTGTACATCAAGAAACCAAGCAAAATGGAAGATT 1080  
Db 1021 ATGATTGCGTAACGAACTTTCTGTACATCAAGAAACCAAGCAAAATGGAAGATT 1080  
QY 1081 TAAATTTCACTGTGGCAGTATGATCACCTCAGGCTACCTATCTTCTGTCTTAGCAAGA 1140  
Db 1081 TAAATTTCACTGTGGCAGTATGATCACCTCAGGCTACCTATCTTCTGTCTTAGCAAGA 1140  
QY 1141 AGGCTCGGGGAAACCAAGTTGTTAAGGCTTTCTTTCTTCTGTGTGAGCAAGCAGTG 1200  
Db 1141 AGGCTCGGGGAAACCAAGTTGTTAAGGCTTTCTTTCTTCTGTGTGAGCAAGCAGTG 1200  
QY 1201 CTACCCCATTCACAGACATCAAGTCAAAATATGGAGTCTGGAAGATGAGCCGCAAGTG 1260  
Db 1201 CTACCCCATTCACAGACATCAAGTCAAAATATGGAGTCTGGAAGATGAGCCGCAAGTG 1260  
QY 1261 ATAAAAATTTATGTGCGGGGATTAATAGACTATGATTTGGTGTGAAGATGATTATCAACAG 1320  
Db 1261 ATAAAAATTTATGTGCGGGGATTAATAGACTATGATTTGGTGTGAAGATGATTATCAACAG 1320  
QY 1321 GTGCTGTACTCTCCCGAATCAAGTTTACAAGTACTGAGCAAGTCAAAATGGGTGG 1380  
Db 1321 GTGCTGTACTCTCCCGAATCAAGTTTACAAGTACTGAGCAAGTCAAAATGGGTGG 1380  
QY 1381 AATCTAAATCAATTAATCCAGCTTATGAGACACCTGCAACCTGCAATTAATTAAGAACAG 1440  
Db 1381 AATCTAAATCAATTAATCCAGCTTATGAGACACCTGCAACCTGCAATTAATTAAGAACAG 1440

1441 AAAATGTATATCTCCTAAGTCTGCTGTAAGAATGAAGTACTTTATGTTTCTGAGC 1500  
1441 AAAATGTATATCTCCTAAGTCTGCTGTAAGAATGAAGTACTTTATGTTTCTGAGC 1500  
1501 CAAAGACTCCAGTAAATAGAACACAGCAATAAGAGAGAAATCTCACTACGCCAAATCGTT 1560  
1501 CAAAGACTCCAGTAAATAGAACACAGCAATAAGAGAGAAATCTCACTACGCCAAATCGTT 1560  
1561 ACATACACCCCTCAAAAGCTAGAAACAGTCCCTGAAAGAACTCCAATTAATAACCAAG 1620  
1561 ACATACACCCCTCAAAAGCTAGAAACAGTCCCTGAAAGAACTCCAATTAATAACCAAG 1620  
1621 TAAATTCACAGGAACAGCAAGTCTAATGACAGGTGTCAATAGCCCTGAGAGCGGTGCC 1680  
1621 TAAATTCACAGGAACAGCAAGTCTAATGACAGGTGTCAATAGCCCTGAGAGCGGTGCC 1680  
1681 GCTCAGTGAATTTGGATCTCAACCAAGCACATATGAGGAGACTCCAAAAGAAAGGGAG 1740  
1681 GCTCAGTGAATTTGGATCTCAACCAAGCACATATGAGGAGACTCCAAAAGAAAGGGAG 1740  
1741 CCAAAGTGTGTGGAGCCTTGAAGGGGTGGATAAGGTATCACTGTGCTCACCAGGA 1800  
1741 CCAAAGTGTGTGGAGCCTTGAAGGGGTGGATAAGGTATCACTGTGCTCACCAGGA 1800  
1801 GCAAAAGGAAGGCTTCTGCCAGAGACGGGCCCAAGAGACTTAAAGCTTCTACTATAATGTGA 1860  
1801 GCAAAAGGAAGGCTTCTGCCAGAGACGGGCCCAAGAGACTTAAAGCTTCTACTATAATGTGA 1860  
1861 CTACAACTAGATTTAGTGAATCCAGATCAACTGTGTGAATGAATAATGTCTATTTCCAA 1920  
1861 CTACAACTAGATTTAGTGAATCCAGATCAACTGTGTGAATGAATAATGTCTATTTCCAA 1920  
1921 AGNAGCATGTGTGCTTGTACAAAAGGCTTATACACTGAAGTGTCAAAACACAGTCAGATT 1980  
1921 AGAAGCATGTGTGCTTGTGTAACAAAAGGCTTATACACTGAAGTGTCAAAACACAGTCAGATT 1980  
1981 TTGGGAAAGTGCAATGCAATTTGAATTAAGAGTGTCCAGCTTCAAAAACCGGATGTGG 2040  
1981 TTGGGAAAGTGCAATGCAATTTGAATTAAGAGTGTCCAGCTTCAAAAACCGGATGTGG 2040  
2041 TGCGTATCAGGAGCGCGCTTAAGGCGATGCGCTGGGTTTCAAAAAGATTAGTGGAA 2100  
2041 TGCGTATCAGGAGCGCGCTTAAGGCGATGCGCTGGGTTTCAAAAAGATTAGTGGAA 2100  
2101 ACATCCTATCTAGCTGCAAGGTATAATTAAGTGAATTTCCATCCTGCCGGATGAGTGTG 2160  
2101 ACATCCTATCTAGCTGCAAGGTATAATTAAGTGAATTTCCATCCTGCCGGATGAGTGTG 2160  
2161 GGTGTGATACAGCTTACATAAGACTGTTATGATCGCTTTGATTTTAAAGTTCATTGGAA 2220  
2161 GGTGTGATACAGCTTACATAAGACTGTTATGATCGCTTTGATTTTAAAGTTCATTGGAA 2220  
2221 CTACCAACTTTGTTTCTAAAGAGCTATCTTAAAGCAATATCTCTTTGTTTTTAAACAAA 2280  
2221 CTACCAACTTTGTTTCTAAAGAGCTATCTTAAAGCAATATCTCTTTGTTTTTAAACAAA 2280  
2281 GATATTTATTTGTGTATGAATCTAAATCAAGCCCATCTGTCAATATGTTTACTGTCTTTT 2340  
2281 GATATTTATTTGTGTATGAATCTAAATCAAGCCCATCTGTCAATATGTTTACTGTCTTTT 2340  
2341 TAATCATGTGTTTTGTATATTAATTAATTTGTTGACCTTTCTAGATTCACTTCATATGTG 2400  
2341 TAATCATGTGTTTTGTATATTAATTAATTTGTTGACCTTTCTAGATTCACTTCATATGTG 2400  
2401 AATGTAAGCTTTAAGCTATGCTCTTTGTAAGTGTAAATTTCTTTCTGAAATAAAACCAT 2460  
2401 AATGTAAGCTTTAAGCTATGCTCTTTGTAAGTGTAAATTTCTTTCTGAAATAAAACCAT 2460  
2461 TTGTGAATAT 2470  
2461 TTGTGAATAT 2470

RESULT 3  
US-09-880-107-1725  
; Sequence 1725, Application US/09880107  
; Patent No. US20020142981A1  
; GENERAL INFORMATION:  
; APPLICANT: Horne, Darci T.  
; APPLICANT: Vockley, Joseph G.  
; APPLICANT: Scherf, Uwe  
; APPLICANT: Gene Logic, Inc.  
; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer  
; FILE REFERENCE: 44921-5028-WO  
; CURRENT APPLICATION NUMBER: US/09/880,107  
; CURRENT FILING DATE: 2001-06-14  
; PRIOR APPLICATION NUMBER: US 60/211,379  
; PRIOR FILING DATE: 2000-06-14  
; PRIOR APPLICATION NUMBER: US 60/237,054  
; PRIOR FILING DATE: 2000-10-02  
; NUMBER OF SEQ ID NOS: 3950  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 1725  
; LENGTH: 2470  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: Genbank Accession No. US20020142981A1 D79997  
US-09-880-107-1725

Query Match 100.0%; Score 2470; DB 3; Length 2470;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2470; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTGGCGGGCGGAAGCGGCCACAAACCGCGCATCGAAAGATTCTTAGGAACCGCGTACCA 60  
DB 1 TTGGCGGGCGGAGCGGCCACAAACCGCGCATCGAAAGATTCTTAGGAACCGCGTACCA 60  
QY 61 GCCGGTCTCTCAGACAGCAGCGCCCTGTCTTCTGCGGCGCGCTCAGCGTGCC 120  
DB 61 GCCGGTCTCTCAGACAGCAGCGCCCTGTCTTCTGTCGGCGCGCTCAGCGTGCC 120  
QY 121 TCCGCCCTCAGGTTCTTTTCTAATTCNAATTAACCTGCAAGAGGACTATGAAGATT 180  
DB 121 TCCGCCCTCAGGTTCTTTTCTAATTCNAATTAACCTGCAAGAGGACTATGAAGATT 180  
QY 181 ATGATGAATCTCTCAATATTAATGAATACATGAATATTTGGGACAGGTGCTTTGCAA 240  
DB 181 ATGATGAATCTCTCAATATTAATGAATACATGAATATTTGGGACAGGTGCTTTGCAA 240  
QY 241 AGGTCAAACTTGCCTGCCATATCTTACTGGAGAGATGGTAGCTATAAAATCATGGATA 300  
DB 241 AGGTCAAACTTGCCTGCCATATCTTACTGGAGAGATGGTAGCTATAAAATCATGGATA 300  
QY 301 AAAACACACTAGGAGTGAATTTGCCCGGATCAAAACCGAGATTGAGGCTTTGAAGAACC 360  
DB 301 AAAACACACTAGGAGTGAATTTGCCCGGATCAAAACCGAGATTGAGGCTTTGAAGAACC 360  
QY 361 TCAGACATCAGCATATATGTCAACTCTACATGTCTAGAGACAGCAACAAATATTCA 420  
DB 361 TCAGACATCAGCATATATGTCAACTCTACATGTCTAGAGACAGCAACAAATATTCA 420  
QY 421 TGGTTCTTGAGTACTGCCCTGGAGGAGCTGTTTGAATATAATTTCCAGGATCGCC 480  
DB 421 TGGTTCTTGAGTACTGCCCTGGAGGAGCTGTTTGAATATAATTTCCAGGATCGCC 480  
QY 481 TGTCAAGAGAGACCGCGGTTGTTCCGTCAGATAGTATCTGCTGTTGCTTATGTCG 540  
DB 481 TGTCAAGAGAGACCGCGGTTGTTCCGTCAGATAGTATCTGCTGTTGCTTATGTCG 540  
QY 541 ACAGCAGGCTTATGCTCACAGGACCTCAAGCCAGAAAATTTGCTGTTGATGATATC 600  
DB 541 ACAGCAGGCTTATGCTCACAGGACCTCAAGCCAGAAAATTTGCTGTTGATGATATC 600  
QY 601 ATAAATTAAGCTGATTTGATCTTTGTTGCTCTGTGCAAAACCCCAAGGGTAAAGGATTACC 660



;/ TITLE OF INVENTION: 8990, 2100, 9288, 64698, 10480, 20893, 33230, 1586, 9943,  
;/ TITLE OF INVENTION: 16334, 68862, 9011, 14031, 6178, 21225, 1420, 32236, 2099,  
;/ TITLE OF INVENTION: 2150, 26583, 2784, 8941, 9811, 27444, 50566 OR 66428 MOLECULES  
;/ FILE REFERENCE: MPI02-020PIRNOWINM  
;/ CURRENT APPLICATION NUMBER: US/10/354,358  
;/ PRIOR FILING DATE: 2003-01-30  
;/ PRIOR APPLICATION NUMBER: US 60/353,600  
;/ PRIOR FILING DATE: 2002-01-31  
;/ PRIOR APPLICATION NUMBER: US 60/364,517  
;/ PRIOR FILING DATE: 2002-03-15  
;/ PRIOR APPLICATION NUMBER: US 60/371,075  
;/ PRIOR FILING DATE: 2002-04-09  
;/ PRIOR APPLICATION NUMBER: US 60/371,507  
;/ PRIOR FILING DATE: 2002-04-10  
;/ PRIOR APPLICATION NUMBER: US 60/372,984  
;/ PRIOR FILING DATE: 2002-04-16  
;/ PRIOR APPLICATION NUMBER: US 60/374,194  
;/ PRIOR FILING DATE: 2002-04-19  
;/ PRIOR APPLICATION NUMBER: US 60/382,995  
;/ PRIOR FILING DATE: 2002-05-24  
;/ PRIOR APPLICATION NUMBER: US 60/385,023  
;/ PRIOR FILING DATE: 2002-05-31  
;/ PRIOR APPLICATION NUMBER: US 60/388,853  
;/ PRIOR FILING DATE: 2002-06-14  
;/ PRIOR APPLICATION NUMBER: US 60/389,395  
;/ PRIOR FILING DATE: 2002-06-17  
;/ Remaining Prior Application data removed - See File Wrapper or PALM.  
;/ NUMBER OF SEQ ID NOS: 122  
;/ SOFTWARE: FastSeq for Windows Version 4.0  
;/ SEQ ID NO 7  
;/ LENGTH: 2470  
;/ TYPE: DNA  
;/ ORGANISM: Homo sapiens  
;/ FEATURE:  
;/ NAME/KEY: CDS  
;/ LOCATION: (171)...(2126)  
;/ US-10-354-358-7

Query Match 100.0%; Score 2470; DB 7; Length 2470;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2470; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 TTGGCGGGGGAAGCGGCCACACCGCGGATCGAAAGATTCTTAGGAACGCGTACCA 60  
DB 1 TTGGCGGGGGAAGCGGCCACACCGCGGATCGAAAGATTCTTAGGAACGCGTACCA 60  
QY 61 GCGCGTCTCTCAGACAGAGCGCCCTGTCCTTCTGTGCGGCGCGCTCAGCCGTGCC 120  
DB 61 GCGCGTCTCTCAGACAGAGCGCCCTGTCCTTCTGTGCGGCGCGCTCAGCCGTGCC 120  
QY 121 TCCGCCCTCAGGTTCCTTTTCTTAATTCCAAATAAACTTGCAGAGAGACTATGAAAGATT 180  
DB 121 TCCGCCCTCAGGTTCCTTTTCTTAATTCCAAATAAACTTGCAGAGAGACTATGAAAGATT 180  
QY 181 ATGATGAACCTTCTCAATATATTGAATATACATGAATATTTGGACAGGTGGCTTTGCAA 240  
DB 181 ATGATGAACCTTCTCAATATATTGAATATACATGAATATTTGGACAGGTGGCTTTGCAA 240  
QY 241 AGGTCAAACTTGGCTGCCATATCTTACTTGGAGAGATGGTAGCTATAAAATCATGGATA 300  
DB 241 AGGTCAAACTTGGCTGCCATATCTTACTTGGAGAGATGGTAGCTATAAAATCATGGATA 300  
QY 301 AAAACACACTAGGAGGTGATTTGCCCGGATCAAAACGGAGATTGAGGCCCTTCAAGAAC 360  
DB 301 AAAACACACTAGGAGGTGATTTGCCCGGATCAAAACGGAGATTGAGGCCCTTCAAGAAC 360  
QY 361 TGAGACATCAGCATATATGTCAACTCTACCATGTGCTTAGACAGAGCCCAAAATATTC 420  
DB 361 TGAGACATCAGCATATATGTCAACTCTACCATGTGCTTAGACAGAGCCCAAAATATTC 420  
QY 421 TGGTCTTGAGTACTGCCCTGGAGGAGAGCTGTTGACTATATAATTTCCAGGATCGCC 480  
DB 421 TGGTCTTGAGTACTGCCCTGGAGGAGAGCTGTTGACTATATAATTTCCAGGATCGCC 480

QY 481 TGTCAAGAGAGAGACACCGGGTGTCTTCCGTGATAGTATCTGCTGTGTCTTATGTGC 540  
DB 481 TGTCAAGAGAGAGACACCGGGTGTCTTCCGTGATAGTATCTGCTGTGTCTTATGTGC 540  
QY 541 ACAGCCAGGGCTATGCTCAGAGGACCTCAAGCCAGAAAAATTTGCTGTGTGATGATC 600  
DB 541 ACAGCCAGGGCTATGCTCAGAGGACCTCAAGCCAGAAAAATTTGCTGTGTGATGATC 600  
QY 601 ATAAATTAAGCTGATTCGACTTTTGTCTGTGCAAAACCCCAAGGTAAACAGGATTACC 660  
DB 601 ATAAATTAAGCTGATTCGACTTTTGTCTGTGCAAAACCCCAAGGTAAACAGGATTACC 660  
QY 661 ATCTACAGACATGCTGTGGGAGTCTGGCTTATGACGACCTGAGTTAATACAGGCAAT 720  
DB 661 ATCTACAGACATGCTGTGGGAGTCTGGCTTATGACGACCTGAGTTAATACAGGCAAT 720  
QY 721 CATATCTTGGATCAGAGGAGATGTTGGAGATGGGATCTGTTATATGTTCTTATGT 780  
DB 721 CATATCTTGGATCAGAGGAGATGTTGGAGATGGGATCTGTTATATGTTCTTATGT 780  
QY 781 GTGGATTTCTACCATTTGATGATGATTAATGATGCTTTATACAAAGAGATTATGAGAG 840  
DB 781 GTGGATTTCTACCATTTGATGATGATTAATGATGCTTTATACAAAGAGATTATGAGAG 840  
QY 841 GAAAAATATGATGTTCCCAAGTGGCTCTCTCCAGTAGCATCTGCTTCTTCAACAAATGC 900  
DB 841 GAAAAATATGATGTTCCCAAGTGGCTCTCTCCAGTAGCATCTGCTTCTTCAACAAATGC 900  
QY 901 TGCAGGTGGACCCAAAGAAACGGATTCTATGAAAAATCTATTGAACCATCCTCGATCA 960  
DB 901 TGCAGGTGGACCCAAAGAAACGGATTCTATGAAAAATCTATTGAACCATCCTCGATCA 960  
QY 961 TGCAGGATTACAACTATCTCTGTGAGTGCAAGCAAGAAATCTTTATACCTCGATG 1020  
DB 961 TGCAGGATTACAACTATCTCTGTGAGTGCAAGCAAGAAATCTTTATACCTCGATG 1020  
QY 1021 ATGATTGGGTAAACAGAACTTTCTGTACATCAGAAAAACAGGCAAAACAAATGGAGATT 1080  
DB 1021 ATGATTGGGTAAACAGAACTTTCTGTACATCAGAAAAACAGGCAAAACAAATGGAGATT 1080  
QY 1081 TAAATTTACTGTGGAGATGATGATCACTCAGCGGTACCTATCTTCTGCTTCTAGCCAGA 1140  
DB 1081 TAAATTTACTGTGGAGATGATGATCACTCAGCGGTACCTATCTTCTGCTTCTAGCCAGA 1140  
QY 1141 AGGCTCGGGGAAAACAGTTCTGTTTAAAGCTTTCTTCTTCTCTGTGCAAGCCAGTG 1200  
DB 1141 AGGCTCGGGGAAAACAGTTCTGTTTAAAGCTTTCTTCTTCTCTGTGCAAGCCAGTG 1200  
QY 1201 CTACCCCATTTCAAGACATCAAGTCAAATTAATTTGGAGTCTGGAAGATGTGACCGCAAGTG 1260  
DB 1201 CTACCCCATTTCAAGACATCAAGTCAAATTAATTTGGAGTCTGGAAGATGTGACCGCAAGTG 1260  
QY 1261 ATAAAAATTTATGTGGGGATTAATAGACTATGATGTTGGTGTGAAGTGAATTTATCAACAG 1320  
DB 1261 ATAAAAATTTATGTGGGGATTAATAGACTATGATGTTGGTGTGAAGTGAATTTATCAACAG 1320  
QY 1321 GTGCTGTACTCCCGAACATCAAGTTTACCAAGTACTGGAAGATGATGATGATGGGGTGG 1380  
DB 1321 GTGCTGTACTCCCGAACATCAAGTTTACCAAGTACTGGAAGATGATGATGATGGGGTGG 1380  
QY 1381 AATCTAAATTCATTAATCCAGCTTATGAGAACCTCAGAACCTCCTCAAAATAAATTAAGAACAAAG 1440  
DB 1381 AATCTAAATTCATTAATCCAGCTTATGAGAACCTCAGAACCTCCTCAAAATAAATTAAGAACAAAG 1440  
QY 1441 AAAATGTATATCTCTTAAGTCTGCTGTAAAGATGAAGAGTACTTTATGTTTCTTGAGC 1500  
DB 1441 AAAATGTATATCTCTTAAGTCTGCTGTAAAGATGAAGAGTACTTTATGTTTCTTGAGC 1500  
QY 1501 CAAAGACTCCAGTTAATAGAACCCAGCATAGAGAGAAATACTCTACTAGCCCAAAATCGTT 1560  
DB 1501 CAAAGACTCCAGTTAATAGAACCCAGCATAGAGAGAAATACTCTACTAGCCCAAAATCGTT 1560

```
QY 1561 AACTACACCCCTCAAAAGCTAGAAACCAAGTGCCTGAAAGAAACCTCCAAATTAATAATACCAG 1620
Db 1561 AACTACACCCCTCAAAAGCTAGAAACCAAGTGCCTGAAAGAAACCTCCAAATTAATAATACCAG 1620
QY 1621 TAAATTCACACGAAACAGACAAAGTAAATGACAGGTGCTATTAGCCTGAGAGCGGTGCC 1680
Db 1621 TAAATTCACACGAAACAGACAAAGTAAATGACAGGTGCTATTAGCCTGAGAGCGGTGCC 1680
QY 1681 GCTCAGTGGAAATTTGGATCTCAACACAGCACATATGAGGAGAGCTCCAAAAGAAAGGGAG 1740
Db 1681 GCTCAGTGGAAATTTGGATCTCAACACAGCACATATGAGGAGAGCTCCAAAAGAAAGGGAG 1740
QY 1741 CCAAAGTGTGAGGAGCTTGAAGGGGTTGGATAGGTATCATCTGCTCACCAGGA 1800
Db 1741 CCAAAGTGTGAGGAGCTTGAAGGGGTTGGATAGGTATCATCTGCTCACCAGGA 1800
QY 1801 GCAAAAGGAGGCTTCTGCCAGAGACGGGCCAGAGACTAAAGCTTCACTATAATGTGA 1860
Db 1801 GCAAAAGGAGGCTTCTGCCAGAGACGGGCCAGAGACTAAAGCTTCACTATAATGTGA 1860
QY 1861 CTACAACTAGATTAGTGAATCCAGATCAACTGTTGAATGAATAATGTCTATTCTTCCAA 1920
Db 1861 CTACAACTAGATTAGTGAATCCAGATCAACTGTTGAATGAATAATGTCTATTCTTCCAA 1920
QY 1921 AGAAGCATGTTGACTTTGTACAAAGGGTTTATACACTGAAGTGTCAACACAGTCAAGTT 1980
Db 1921 AGAAGCATGTTGACTTTGTACAAAGGGTTTATACACTGAAGTGTCAACACAGTCAAGTT 1980
QY 1981 TTGGGAAAGTGACAAATGCAATTTGAATTAAGAGTGTCCAGCTTCAAAAACCCGATGTGG 2040
Db 1981 TTGGGAAAGTGACAAATGCAATTTGAATTAAGAGTGTCCAGCTTCAAAAACCCGATGTGG 2040
QY 2041 TGGGTATCAGGAGCGAGCGCTTAAGGCGATGCTGGGTTTACAAAAGATTAGTGGAA 2100
Db 2041 TGGGTATCAGGAGCGAGCGCTTAAGGCGATGCTGGGTTTACAAAAGATTAGTGGAA 2100
QY 2101 ACATCCTATCTAGCTCAAGGTATAATGATGATGATTCCTCCATCCTGCCGAGTGTG 2160
Db 2101 ACATCCTATCTAGCTCAAGGTATAATGATGATGATTCCTCCATCCTGCCGAGTGTG 2160
QY 2161 GGTGTGATACAGCTACATAAGAGCTGTTATGATCGCTTTGATTTTAAAGTTCAATTGGAA 2220
Db 2161 GGTGTGATACAGCTACATAAGAGCTGTTATGATCGCTTTGATTTTAAAGTTCAATTGGAA 2220
QY 2221 CTACCAACTGTTTCTTAAAGAGCTATCTTAAGACCAATATCTCTTTGTTTTTAAACAAA 2280
Db 2221 CTACCAACTGTTTCTTAAAGAGCTATCTTAAGACCAATATCTCTTTGTTTTTAAACAAA 2280
QY 2281 GATATTAATTTTGTGATGATCTAAATCAAGCCCATCTGTCAATATGTTACTGTCTTTT 2340
Db 2281 GATATTAATTTTGTGATGATCTAAATCAAGCCCATCTGTCAATATGTTACTGTCTTTT 2340
QY 2341 TAATCATGTGTTTGTATTAATTAATTTGTTGACTTTCTTGAATGTTGTAATTAATTAAT 2400
Db 2341 TAATCATGTGTTTGTATTAATTAATTTGTTGACTTTCTTGAATGTTGTAATTAATTAAT 2400
QY 2401 AATGTAAGCTCTTAATCATGTCTCTTTGTAATGTTGTAATTTCTTCTGAAATAAACCAT 2460
Db 2401 AATGTAAGCTCTTAATCATGTCTCTTTGTAATGTTGTAATTTCTTCTGAAATAAACCAT 2460
QY 2461 TTGTGAATAT 2470
Db 2461 TTGTGAATAT 2470
```

## RESULT 5

```
US-10-172-118-1535
; Sequence 1535, Application US/10172118
; Publication NO. US20030224374A1
; GENERAL INFORMATION:
; APPLICANT: Dai, Hongyue
; APPLICANT: He, Yudong
; APPLICANT: Linsley, Peter
```

```
; APPLICANT: Mao, Mao
; APPLICANT: Roberts, Chris
; APPLICANT: Van 't Veer, Laura
; APPLICANT: Van de Vijver, Marc
; APPLICANT: Bernards, Rene
; TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
; FILE REFERENCE: 9301-1175-999
; CURRENT APPLICATION NUMBER: US/10/172,118
; PRIOR FILING DATE: 2002-06-14
; PRIOR APPLICATION NUMBER: 60/380,770
; NUMBER OF SEQ ID NOS: 2699
; SEQ ID NO 1535
; LENGTH: 2470
; TYPE: DNA
; ORGANISM: Homo sapiens
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: NM 014791
; DATABASE ENTRY DATE: 2001-06-18
US-10-172-118-1535
```

Query Match 100.0%; Score 2470; DB 7; Length 2470;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 2470; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 TTGGCGGGCGGAAGCGGCCACAAACCCGCGATCGAAAAAGATTCTTTAGGAACGCCCGTACCA 60
Db 1 TTGGCGGGCGGAAGCGGCCACAAACCCGCGATCGAAAAAGATTCTTTAGGAACGCCCGTACCA 60
QY 61 GCCGCTCTCTCAGACAGCAGCGCCCTGTCTTCTGTGGCGCGCTCAGCGTGC 120
Db 61 GCCGCTCTCTCAGACAGCAGCGCCCTGTCTTCTGTGGCGCGCTCAGCGTGC 120
QY 121 TCCGCCCTCAGGTTCTTTTCTTAATTCCAAATAAATTTGCAAGAGGACTATGAAGATT 180
Db 121 TCCGCCCTCAGGTTCTTTTCTTAATTCCAAATAAATTTGCAAGAGGACTATGAAGATT 180
QY 181 ATGATGAATCTCTCAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 240
Db 181 ATGATGAATCTCTCAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 240
QY 241 AGGTCAAACTTGCCTGCCATATCTTACTTGGAGAGATGTTAGTATATAAATCATGGATA 300
Db 241 AGGTCAAACTTGCCTGCCATATCTTACTTGGAGAGATGTTAGTATATAAATCATGGATA 300
QY 301 AAAACACACTAGGAGTGAATTTGCCCGCGATCAAAACCGAGATTCAGGCTTTGAAGAACC 360
Db 301 AAAACACACTAGGAGTGAATTTGCCCGCGATCAAAACCGAGATTCAGGCTTTGAAGAACC 360
QY 361 TGAGACATCAGCATATATGTCAACTCTACCTTACCTAGAGACAGCCCAAAATATTCA 420
Db 361 TGAGACATCAGCATATATGTCAACTCTACCTTACCTAGAGACAGCCCAAAATATTCA 420
QY 421 TGGTCTTGTAGTACTGCCCTGGAGGAGCTGTTGACTATATAATTTCCAGGATCGCC 480
Db 421 TGGTCTTGTAGTACTGCCCTGGAGGAGCTGTTGACTATATAATTTCCAGGATCGCC 480
QY 481 TGTCAAGAGAGAGACCGGGTTGTTCCGTCAGATAGTATCTGCTGTGCTTATGTC 540
Db 481 TGTCAAGAGAGAGACCGGGTTGTTCCGTCAGATAGTATCTGCTGTGCTTATGTC 540
QY 541 ACAGCCAGGCTATGCTCAGAGGACCTCAAGCCAGAAAAATTTGCTGTTGATGAATATC 600
Db 541 ACAGCCAGGCTATGCTCAGAGGACCTCAAGCCAGAAAAATTTGCTGTTGATGAATATC 600
QY 601 ATAAATTAAGCTGATTGACTTTGTCTCTGTGCAAAACCCCAAGGTAACAGGATTACC 660
Db 601 ATAAATTAAGCTGATTGACTTTGTCTCTGTGCAAAACCCCAAGGTAACAGGATTACC 660
QY 661 ATCTACAGACATGCTGTGGAGTCTGGCTTATGACGACCTGAGTTAATACAAGCAAT 720
Db 661 ATCTACAGACATGCTGTGGAGTCTGGCTTATGACGACCTGAGTTAATACAAGCAAT 720
```

QY 721 CATATCTTGGATCAGAGGAGATGTTTGGAGCATGGGCATCTGTTATATGTTCTTATGT 780  
DB 721 CATATCTTGGATCAGAGGAGATGTTTGGAGCATGGGCATCTGTTATATGTTCTTATGT 780  
QY 781 GTGGATTTCTACATTTGATGATGAATAATGTAATGGCTTTATACAAGAGATTTATGAGAG 840  
DB 781 GTGGATTTCTACCATTTGATGATGAATAATGTAATGGCTTTATACAAGAGATTTATGAGAG 840  
QY 841 GAAAAATATGATGTTCCCAAGTGCTCTCTCCAGTAGCATCTCGCTTCTTCAACAATGC 900  
DB 841 GAAAAATATGATGTTCCCAAGTGCTCTCTCCAGTAGCATCTCGCTTCTTCAACAATGC 900  
QY 901 TGCAGGTGGACCCCAAGAAACGGATTTCTATGAAAAATCTATTGAACCATCCCTGGATCA 960  
DB 901 TGCAGGTGGACCCCAAGAAACGGATTTCTATGAAAAATCTATTGAACCATCCCTGGATCA 960  
QY 961 TGCAGGATTAACAATCTATCTCTGTTGAGTGGCAAGCAAGAAATCCTTTTATTCACCTCGATG 1020  
DB 961 TGCAGGATTAACAATCTATCTCTGTTGAGTGGCAAGCAAGAAATCCTTTTATTCACCTCGATG 1020  
QY 1021 ATGATTTGCCGTAAACAGAACTTCTGTACATCACAGAAACACAGGCAAAACAATGGAGATT 1080  
DB 1021 ATGATTTGCCGTAAACAGAACTTCTGTACATCACAGAAACACAGGCAAAACAATGGAGATT 1080  
QY 1081 TAAATTTCACTGTGGCAGTATGATCACTCACGGCTACCTATCTCTGCTTTCTAGCCAAGA 1140  
DB 1081 TAAATTTCACTGTGGCAGTATGATCACTCACGGCTACCTATCTCTGCTTTCTAGCCAAGA 1140  
QY 1141 AGGCTCGGGAAAAACGATTTGCTTTAAGGCTTTCTTTCTCTCTGTGGCAAGCCAGTG 1200  
DB 1141 AGGCTCGGGAAAAACGATTTGCTTTAAGGCTTTCTTTCTCTCTGTGGCAAGCCAGTG 1200  
QY 1201 CTACCCCATTTACAGACATCAAGTCAAAATTTGGAGTCTGGAAGATGTGACCGCAAGTG 1260  
DB 1201 CTACCCCATTTACAGACATCAAGTCAAAATTTGGAGTCTGGAAGATGTGACCGCAAGTG 1260  
QY 1261 ATAAAAATATGTGGGGGATTAATAGACTATGATTTGGTGTGAAGATGATTTATCAACAG 1320  
DB 1261 ATAAAAATATGTGGGGGATTAATAGACTATGATTTGGTGTGAAGATGATTTATCAACAG 1320  
QY 1321 GTCTGCTACTCCCGCAACATCAAGTTTACCAAGTACTGGAAGATCAAAATCAAAATGGGGTG 1380  
DB 1321 GTCTGCTACTCCCGCAACATCAAGTTTACCAAGTACTGGAAGATCAAAATGGGGTG 1380  
QY 1381 AATCTAAATCAATTAATCTCCAGCTTATGCAAGACACCTGCAAAATTAATTAAGAAACAAG 1440  
DB 1381 AATCTAAATCAATTAATCTCCAGCTTATGCAAGACACCTGCAAAATTAATTAAGAAACAAG 1440  
QY 1441 AAAATGTATATACTCTCTAAGTCTGCTGTAAGAAATGAAGATGATTTATGTTTCTCTGAGC 1500  
DB 1441 AAAATGTATATACTCTCTAAGTCTGCTGTAAGAAATGAAGATGATTTATGTTTCTCTGAGC 1500  
QY 1501 CAAAGACTCCAGTTAATAAGAACAGCATTAAGAGAGAAATPACTCATCGCCAAATCGTT 1560  
DB 1501 CAAAGACTCCAGTTAATAAGAACAGCATTAAGAGAGAAATPACTCATCGCCAAATCGTT 1560  
QY 1561 ACATACACCTCAAAAGCTAGAAACAGTGCCTGAAAGAACTCCAATTAATTAATACCAAG 1620  
DB 1561 ACATACACCTCAAAAGCTAGAAACAGTGCCTGAAAGAACTCCAATTAATTAATACCAAG 1620  
QY 1621 TAAATCAACAGGAACAGACAAAGTTTAATGACAGTGTCAATTAGCCCTGAGAGCGGTGCC 1680  
DB 1621 TAAATCAACAGGAACAGACAAAGTTTAATGACAGTGTCAATTAGCCCTGAGAGCGGTGCC 1680  
QY 1681 GCTCAGTGGAAATGGATCTCAACCAAGCAATATGAGGAGACTCCAAAAAGAAAGGAG 1740  
DB 1681 GCTCAGTGGAAATGGATCTCAACCAAGCAATATGAGGAGACTCCAAAAAGAAAGGAG 1740  
QY 1741 CCAAAGTGTGGAGCGCTTGAAGGGGTGGATAGGTTATCACTGTCTCACCCAGGA 1800  
DB 1741 CCAAAGTGTGGAGCGCTTGAAGGGGTGGATAGGTTATCACTGTCTCACCCAGGA 1800  
QY 1801 GCAAAAGGAAGGGTTCTGCCAGAGACGGGGCCAGAGACTAAAGCTTCACTATAATGTGA 1860

DB 1801 GCAAAAGGAAGGGTTCTGCCAGAGACGGGGCCAGAGACTAAAGCTTCACTATAATGTGA 1860  
QY 1861 CTACAACTAGATTAGTGAATCCAGATCAACTGTTCAATGAATAATGTCATTTCTTCCAA 1920  
DB 1861 CTACAACTAGATTAGTGAATCCAGATCAACTGTTGAATGAATAATGTCATTTCTTCCAA 1920  
QY 1921 AGAAGCATGTTGACTTTGTACAAAAGGGTTATACACTGAAAGTGTCAAAACACAGTCAGATT 1980  
DB 1921 AGAAGCATGTTGACTTTGTACAAAAGGGTTATACACTGAAAGTGTCAAAACACAGTCAGATT 1980  
QY 1981 TTGGGAAAGTGACAAATGCAATTTGAAATTAAGAGTGTGCCAGCTTCAAAAACCCGATGG 2040  
DB 1981 TTGGGAAAGTGACAAATGCAATTTGAAATTAAGAGTGTGCCAGCTTCAAAAACCCGATGG 2040  
QY 2041 TGGGTATCAGGAGGAGCGGCTTAAAGGGGATGCGCTTGAAGAGATTTAGTGGAG 2100  
DB 2041 TGGGTATCAGGAGGAGCGGCTTAAAGGGGATGCGCTTGAAGAGATTTAGTGGAG 2100  
QY 2101 ACATCTATCTAGCTGCAAGGTATAATGATGGATTTCTTCCATCCTGCCGGATGAGTGTG 2160  
DB 2101 ACATCTATCTAGCTGCAAGGTATAATGATGGATTTCTTCCATCCTGCCGGATGAGTGTG 2160  
QY 2161 GGTGTGATACAGCTTACATAAAGACTGTTATGATGATCGCTTTGATTTTAAAGTTCATTGGAA 2220  
DB 2161 GGTGTGATACAGCTTACATAAAGACTGTTATGATGATCGCTTTGATTTTAAAGTTCATTGGAA 2220  
QY 2221 CTACCAACTGTTTCTTAAGAGCTATCTTAAGACCAATATCTCTTTGTTTTTAAACAAA 2280  
DB 2221 CTACCAACTGTTTCTTAAGAGCTATCTTAAGACCAATATCTCTTTGTTTTTAAACAAA 2280  
QY 2281 GATATTATTTCTGTATGAATCTAAATCAAGCCCATCTGTCAATATGTTTACTGTCTTTTT 2340  
DB 2281 GATATTATTTCTGTATGAATCTAAATCAAGCCCATCTGTCAATATGTTTACTGTCTTTTT 2340  
QY 2341 TAAATCATGTTGTTGTATATTAATAATTAATTTGTTGACTTTCTTAGATTCACTTCCATATGTG 2400  
DB 2341 TAAATCATGTTGTTGTATATTAATAATTAATTTGTTGACTTTCTTAGATTCACTTCCATATGTG 2400  
QY 2401 AATGTAAAGCTTAACTATGCTCTTCTGTAAGTGTAAATTTCTTCTGAAATAAAACCAT 2460  
DB 2401 AATGTAAAGCTTAACTATGCTCTTCTGTAAGTGTAAATTTCTTCTGAAATAAAACCAT 2460  
QY 2461 TTGTGAATAT 2470  
DB 2461 TTGTGAATAT 2470

RESULT 6

US-10-295-027-505  
; Sequence 505, Application US/10295027  
; Publication No. US20030232350A1  
; GENERAL INFORMATION:  
; APPLICANT: Afar, Daniel  
; APPLICANT: Aziz, Natasha  
; APPLICANT: Ginsberg, Wendy M.  
; APPLICANT: Gish, Kurt C.  
; APPLICANT: Glyme, Richard  
; APPLICANT: Hevezi, Peter A.  
; APPLICANT: Mack, David H.  
; APPLICANT: Murray, Richard  
; APPLICANT: Watson, Susan R.  
; APPLICANT: Eos Biotechnology, Inc.  
; TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and  
; TITLE OF INVENTION: Methods of Screening for Modulators of Cancer  
; FILE REFERENCE: 018501-012500US  
; CURRENT APPLICATION NUMBER: US/10/295,027  
; CURRENT FILING DATE: 2002-11-13  
; PRIOR APPLICATION NUMBER: US 09/663,733  
; PRIOR FILING DATE: 2000-09-15  
; PRIOR APPLICATION NUMBER: US 60/350,666  
; PRIOR FILING DATE: 2001-11-13  
; PRIOR APPLICATION NUMBER: US 60/335,394



```
; PRIOR FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: US 60/332,464
; PRIOR FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: US 60/334,393
; PRIOR FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: US 60/340,376
; PRIOR FILING DATE: 2001-12-14
; PRIOR APPLICATION NUMBER: US 60/347,211
; PRIOR FILING DATE: 2002-01-08
; PRIOR APPLICATION NUMBER: US 60/347,349
; PRIOR FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 60/355,250
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: US 60/356,714
; PRIOR FILING DATE: 2002-02-13
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1386
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 505
; LENGTH: 2470
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-295-027-505

Query Match
Best Local Similarity 100.0%; Score 2470; DB 7; Length 2470;
Matches 2470; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTGGCGGGCGGAAGCGGCCACACCGCGGATCGAAAGATTCTTAGGAACGCGGTACCA 60
DB 1 TTGGCGGGCGGAAGCGGCCACACCGCGGATCGAAAGATTCTTAGGAACGCGCGTACCA 60
QY 61 GCGCGCTCTCTCAGGACAGCAGCGCCCTGTCTTCTGTGCGGGCGCGCTCAGCGGTGCC 120
DB 61 GCGCGCTCTCTCAGGACAGCAGCGCCCTGTCTTCTGTGCGGGCGCGCTCAGCGGTGCC 120
QY 121 TCGGCGCCCTCAGGTTCTTTCTTAATTCAAATAAACTTGCAAGAGACTATGAAGATT 180
DB 121 TCGGCGCCCTCAGGTTCTTTCTTAATTCCAATAAACTTGCAAGAGACTATGAAGATT 180
QY 181 ATGATGAACCTTCTCAATAATATTGAATATACATGAATACTATGGACAGGTGGCTTTGCAA 240
DB 181 ATGATGAACCTTCTCAATAATATTGAATATACATGAATACTATGGACAGGTGGCTTTGCAA 240
QY 241 AGGTCAAACTTGCTGCCATATCTTTACTTGGAGAGATGGTAGCTATAAAATCATGGATA 300
DB 241 AGGTCAAACTTGCTGCCATATCTTTACTTGGAGAGATGGTAGCTATAAAATCATGGATA 300
QY 301 AAAACACACTAGGAGTATTTGCCCGGATCAAAACGGAGATTGAGGCTTTGAAGAACC 360
DB 301 AAAACACACTAGGAGTATTTGCCCGGATCAAAACGGAGATTGAGGCTTTGAAGAACC 360
QY 361 TGAGACATCAGCATATATGTCAACTTACCATGTGTAGACAGCAGCAACAAATATTTCA 420
DB 361 TGAGACATCAGCATATATGTCAACTTACCATGTGTAGACAGCAGCAACAAATATTTCA 420
QY 421 TGGTTCTTGAGTACTGCCCTTGAGGAGAGCTGTGTGACTATATAATTTCCAGGATCGCC 480
DB 421 TGGTTCTTGAGTACTGCCCTTGAGGAGAGCTGTGTGACTATATAATTTCCAGGATCGCC 480
QY 481 TGTCAAGAGAGAGACCGGGTGTCTTCGTCAGATAGTATCTGCTGTGCTTATGTGTC 540
DB 481 TGTCAAGAGAGAGACCGGGTGTCTTCGTCAGATAGTATCTGCTGTGCTTATGTGTC 540
QY 541 ACAGCCAGGGCTATGCTCAGCGGACTCAAGCCAGAAATTTGCTGTGTTGATGATATC 600
DB 541 ACAGCCAGGGCTATGCTCAGCGGACTCAAGCCAGAAATTTGCTGTGTTGATGATATC 600
QY 601 ATAAATTAAGCTGATTGACTTTGGTCTCTGTGCAAAACCCAGGGGTAAACAAGGATTACC 660
DB 601 ATAAATTAAGCTGATTGACTTTGGTCTCTGTGCAAAACCCAGGGGTAAACAAGGATTACC 660
QY 661 ATCTACAGACATGCTGTGGGAGTCTGGCTTATGACGACACCTGAGTTAATACAAGGCAAT 720
```

Db 1741 CCAAAGTGTGGAGCCTTGAAGGGGGTGGATAAGGTTATCACTGTGCTCACCAGGA 1800  
Qy 1801 GCAAAAGGAGGGTCTGCGAGAGACGGGCCAGAGACTAAAGCTTCACTATAATGTGA 1860  
Db 1801 GCAAAAGGAGGGTCTGCGAGAGACGGGCCAGAGACTAAAGCTTCACTATAATGTGA 1860  
Qy 1861 CTACAACTAGATTAGTGAATCCAGATCAACTGTGTAATGAATAATGTCTATTCTCCAA 1920  
Db 1861 CTACAACTAGATTAGTGAATCCAGATCAACTGTGTAATGAATAATGTCTATTCTCCAA 1920  
Qy 1921 AGAAGCATGTGTGATTTGTATCAAAAAGGGTTATACATGAAGTGTCAAAACACAGTCAAGTT 1980  
Db 1921 AGAAGCATGTGTGATTTGTATCAAAAAGGGTTATACATGAAGTGTCAAAACACAGTCAAGTT 1980  
Qy 1981 TTGGGAAAGTGAATGCAATTTGAATTAAGAGTGTGCCAGCTTCAAAACCCGATGTGG 2040  
Db 1981 TTGGGAAAGTGAATGCAATTTGAATTAAGAGTGTGCCAGCTTCAAAACCCGATGTGG 2040  
Qy 2041 TGGGTATCAGGAGCGAGCGCTTAAAGGGGATGCCCTGGGTTTACAAAAGATTAGTGGAA 2100  
Db 2041 TGGGTATCAGGAGCGAGCGCTTAAAGGGGATGCCCTGGGTTTACAAAAGATTAGTGGAA 2100  
Qy 2101 ACATCCTATCTAGCTGCAAGGTATAATTTGATGATTTCTCCATCTGCCGGATGATGTG 2160  
Db 2101 ACATCCTATCTAGCTGCAAGGTATAATTTGATGATTTCTCCATCTGCCGGATGATGTG 2160  
Qy 2161 GGTGTGATACAGCTACATAAAGACTGTTATGATCGCTTTGATTTTAAAGTTCATTGGAA 2220  
Db 2161 GGTGTGATACAGCTACATAAAGACTGTTATGATCGCTTTGATTTTAAAGTTCATTGGAA 2220  
Qy 2221 CTACCAACTTGTCTTAAAGAGCTATCTTAAAGACCAATATCTTTTGTGTTTAAACAAA 2280  
Db 2221 CTACCAACTTGTCTTAAAGAGCTATCTTAAAGACCAATATCTTTTGTGTTTAAACAAA 2280  
Qy 2281 GATATTATTTTGTGATGAATCTAAATCAAGCCCATCTGTCATATGTTACTGTCTTTT 2340  
Db 2281 GATATTATTTTGTGATGAATCTAAATCAAGCCCATCTGTCATATGTTACTGTCTTTT 2340  
Qy 2341 TAATCATGTGTTTGTATATTAATTAATTTGTTGACTTCTTAGATTCACTTCCATATGTG 2400  
Db 2341 TAATCATGTGTTTGTATATTAATTAATTTGTTGACTTCTTAGATTCACTTCCATATGTG 2400  
Qy 2401 AATGTAAGCTCTTAACTATGCTCTTTGTAATGTGTAATTTCTTCTGAAATAAAACCAT 2460  
Db 2401 AATGTAAGCTCTTAACTATGCTCTTTGTAATGTGTAATTTCTTCTGAAATAAAACCAT 2460  
Qy 2461 TTGTGAATAT 2470  
Db 2461 TTGTGAATAT 2470

RESULT 7  
US-10-173-999-126  
; Sequence 126, Application US/10173999  
; Publication No. US2004000563A1  
; GENERAL INFORMATION:  
; APPLICANT: Mack, David H.  
; APPLICANT: Gish, Kurt C.  
; APPLICANT: Eos Biotechnology, Inc.  
; TITLE OF INVENTION: Methods of Diagnosis of Ovarian Cancer, Compositions  
; TITLE OF INVENTION: and Methods of Screening for Modulators of Ovarian  
; TITLE OF INVENTION: Cancer  
; FILE REFERENCE: 018501-002420US  
; CURRENT APPLICATION NUMBER: US/10173,999  
; CURRENT FILING DATE: 2002-06-17  
; PRIOR APPLICATION NUMBER: US 60/299,234  
; PRIOR FILING DATE: 2001-06-18  
; PRIOR APPLICATION NUMBER: US 60/315,287  
; PRIOR FILING DATE: 2001-08-27  
; PRIOR APPLICATION NUMBER: US 60/350,666  
; PRIOR FILING DATE: 2001-11-13  
; PRIOR APPLICATION NUMBER: US 60/372,246  
; PRIOR FILING DATE: 2001-04-12

; NUMBER OF SEQ ID NOS: 163  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 126  
; LENGTH: 2470  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-173-999-126  
  
Query Match 100.0%; Score 2470; DB 7; Length 2470;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2470; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
Qy 1 TTGGCGGGCGGAAGCGGCCACAACCCGGCGATCGAAAGATTCTTTAGGAACCCCGTACCA 60  
Db 1 TTGGCGGGCGGAAGCGGCCACAACCCGGCGATCGAAAGATTCTTTAGGAACCCCGTACCA 60  
Qy 61 GCCGGTCTCTCAGAGACAGAGCCCTGCTCTTCTGTGGCGCGCTCAGCCGTGCC 120  
Db 61 GCCGGTCTCTCAGAGACAGAGCCCTGCTCTTCTGTGGCGCGCTCAGCCGTGCC 120  
Qy 121 TCCGCCCTCAGGTTCTTTTCTAATTCCAAATAAACTTGCAGAGAGACTATGAAGATT 180  
Db 121 TCCGCCCTCAGGTTCTTTTCTAATTCCAAATAAACTTGCAGAGAGACTATGAAGATT 180  
Qy 181 ATGATGAACCTTCTCAAAATATTATGAATTAACATGAATTAATGGGACAGGTGGCTTTGCAA 240  
Db 181 ATGATGAACCTTCTCAAAATATTATGAATTAACATGAATTAATGGGACAGGTGGCTTTGCAA 240  
Qy 241 AGGTCAAACTTGCCTGCCATATCCTTACTGGAGAGATGCTAGCTATAAAATCATGGATA 300  
Db 241 AGGTCAAACTTGCCTGCCATATCCTTACTGGAGAGATGCTAGCTATAAAATCATGGATA 300  
Qy 301 AAAACACACTAGGGAGTGAATTTGCCCGGATCAAAACGGAGATTGAGGCTTTGAAGAACC 360  
Db 301 AAAACACACTAGGGAGTGAATTTGCCCGGATCAAAACGGAGATTGAGGCTTTGAAGAACC 360  
Qy 361 TGAGACATCAGCATATATGTCAACTCTACCATGTCTAGAGACAGCCAAACAAATATTCA 420  
Db 361 TGAGACATCAGCATATATGTCAACTCTACCATGTCTAGAGACAGCCAAACAAATATTCA 420  
Qy 421 TGGTCTTGAGTACTGCCCTGGAGAGAGCTGTTTGACTATATAATTTCCAGGATCGCC 480  
Db 421 TGGTCTTGAGTACTGCCCTGGAGAGAGCTGTTTGACTATATAATTTCCAGGATCGCC 480  
Qy 481 TGTCAAGAGAGAGACCCGGGTGTCTTCCGTCAAGTAGTATCTGCTGTGTTGTTATGTC 540  
Db 481 TGTCAAGAGAGAGACCCGGGTGTCTTCCGTCAAGTAGTATCTGCTGTGTTGTTATGTC 540  
Qy 541 ACAGCCAGGGCTATGCTCAAGGGACCTCAAGCCAGGAAATTTGCTGTTGATGAATATC 600  
Db 541 ACAGCCAGGGCTATGCTCAAGGGACCTCAAGCCAGGAAATTTGCTGTTGATGAATATC 600  
Qy 601 ATAAATTAAGCTGATTGACTTTGGTCTGTGTGCAAAACCCAGGGTAAACAAGGATATCC 660  
Db 601 ATAAATTAAGCTGATTGACTTTGGTCTGTGTGCAAAACCCAGGGTAAACAAGGATATCC 660  
Qy 661 ATCTACAGACATGCTGTGGAGTCTGTGGCTTATGACGACCTCAGTGTAAATCAAGGCAAT 720  
Db 661 ATCTACAGACATGCTGTGGAGTCTGTGGCTTATGACGACCTCAGTGTAAATCAAGGCAAT 720  
Qy 721 CATATCTTGGATCAGAGGAGATGTTTGGAGCATGGGCATCTGTTTATATGTTCTTATGT 780  
Db 721 CATATCTTGGATCAGAGGAGATGTTTGGAGCATGGGCATCTGTTTATATGTTCTTATGT 780  
Qy 781 GTGGATTTCTACATTTGATGATGATGTAATGTAATGGCTTTATACAGAGATATCAGAG 840  
Db 781 GTGGATTTCTACATTTGATGATGATGTAATGTAATGGCTTTATACAGAGATATCAGAG 840  
Qy 841 GAAATATGATGTTTCCCAAGTGGCTCTCTCCAGTAGCATTTCTGCTTCTTCAACAATGC 900  
Db 841 GAAATATGATGTTTCCCAAGTGGCTCTCTCCAGTAGCATTTCTGCTTCTTCAACAATGC 900  
Qy 901 TGCAGGTGACCCAAAGAAACGGATTTTCTATGAAATAATCTATTGAACCATCCCTGGATCA 960

Db 901 |||||TGAGGTGGACCCAAAGAAACGAGATTCTATGAAAAATCTATTGAACCATCCCTGGATCA 960  
QY 961 TGCAGATTACAACATCTCTGTGTGAGTGGCAAAAGCAAGAAATCTTTTATTACCTCGATG 1020  
Db 961 TGCAGATTACAACATCTCTGTGTGAGTGGCAAAAGCAAGAAATCTTTTATTACCTCGATG 1020  
QY 1021 ATGATTGGCGTAAACAGAACTTTCTGTATCATCACAGAAACAAACGCGCAAAACAATGGAGATT 1080  
Db 1021 ATGATTGGCGTAAACAGAACTTTCTGTATCATCACAGAAACAAACGCGCAAAACAATGGAGATT 1080  
QY 1081 TAATTTCACCTGTGGCAGTAGATCACTCACGGCTACCTATCTTCTGTCTGTAGCCAGA 1140  
Db 1081 TAATTTCACCTGTGGCAGTAGATCACTCACGGCTACCTATCTTCTGTCTGTAGCCAGA 1140  
QY 1141 AGCTCGGGGAAAAACAGTTTCGTTTAAGGCTTTCTTCTCTCTGTGGCAAGCCAGTG 1200  
Db 1141 AGCTCGGGGAAAAACAGTTTCGTTTAAGGCTTTCTTCTCTCTGTGGCAAGCCAGTG 1200  
QY 1201 CTACCCCATTCACAGATCAAGTCAAAATTAATTGGAGTCTGGAAAGATGTGACCGCAAGTG 1260  
Db 1201 CTACCCCATTCACAGATCAAGTCAAAATTAATTGGAGTCTGGAAAGATGTGACCGCAAGTG 1260  
QY 1261 ATAAAAATTATGTGGCGGGATTAATAGACTATGATTGGTGTGAAGATGATTATCAACAG 1320  
Db 1261 ATAAAAATTATGTGGCGGGATTAATAGACTATGATTGGTGTGAAGATGATTATCAACAG 1320  
QY 1321 GTGCTGCTACTCCCGGAAACATACAGTTTACCAAGTACTGGACAGAAATCAATGGGGTGG 1380  
Db 1321 GTGCTGCTACTCCCGGAAACATACAGTTTACCAAGTACTGGACAGAAATCAATGGGGTGG 1380  
QY 1381 AATCTAAATCATTAATCCAGCCTTATGCAGAACACCTCGCAATAAATTAAGAAACAAAG 1440  
Db 1381 AATCTAAATCATTAATCCAGCCTTATGCAGAACACCTCGCAATAAATTAAGAAACAAAG 1440  
QY 1441 AAATGTATATATCTCTAAGTCTGCTGTAAAGATCAAGATGACTTATTTCTCTGAGC 1500  
Db 1441 AAATGTATATATCTCTAAGTCTGCTGTAAAGATCAAGATGACTTATTTCTCTGAGC 1500  
QY 1501 CAAAGACTCCAGTTAATAAGAACCCAGCATAAGAGAGAAATCTCACTACGCCAAATCGTT 1560  
Db 1501 CAAAGACTCCAGTTAATAAGAACCCAGCATAAGAGAGAAATCTCACTACGCCAAATCGTT 1560  
QY 1561 ACATACACCCCTCAAAGCTAGAAACCGAGTCTGCTGAAAGAACTCCAAATTAATAACAG 1620  
Db 1561 ACATACACCCCTCAAAGCTAGAAACCGAGTCTGCTGAAAGAACTCCAAATTAATAACAG 1620  
QY 1621 TAAATTCACAGGAACAGACAGATTATGCACAGTGTCTATTAGCCCTGAGAGCGGTGCC 1680  
Db 1621 TAAATTCACAGGAACAGACAGATTATGCACAGTGTCTATTAGCCCTGAGAGCGGTGCC 1680  
QY 1681 GCTCAGTGGAAATGGATCTCAACCAAGCACATATGGAGAGACTCCCAAAAGAAAGGGAG 1740  
Db 1681 GCTCAGTGGAAATGGATCTCAACCAAGCACATATGGAGAGACTCCCAAAAGAAAGGGAG 1740  
QY 1741 CCAAAGTGTGTGGAGCCTTGAAGGGGTGGATAAGGTTATCACTGTGCTCACCAGGA 1800  
Db 1741 CCAAAGTGTGTGGAGCCTTGAAGGGGTGGATAAGGTTATCACTGTGCTCACCAGGA 1800  
QY 1801 GCAAAAGGAAGGGTCTGCGCAGAGACGGCCCGAGAGACTAAAGCTTCACTATAATGTGA 1860  
Db 1801 GCAAAAGGAAGGGTCTGCGCAGAGACGGCCCGAGAGACTAAAGCTTCACTATAATGTGA 1860  
QY 1861 CTACAACTAGATTAGTGAATCCAGATCAACTGTTGAATCAAAATGATCTTCTTCCAA 1920  
Db 1861 CTACAACTAGATTAGTGAATCCAGATCAACTGTTGAATCAAAATGATCTTCTTCCAA 1920  
QY 1921 AGAAGCATGTTGACTTTGTACAAAAGGGTTATACACTGAAGTGTCAAAACACAGTCAGATT 1980  
Db 1921 AGAAGCATGTTGACTTTGTACAAAAGGGTTATACACTGAAGTGTCAAAACACAGTCAGATT 1980  
QY 1981 TTGGGAAAGTGAACATGCAATTTGAATTAGAGTGTGCCAGCTTCAAAAACCCGATGTGG 2040  
|||||

Db 1981 TTGGGAAAGTGAACATGCAATTGGAATTAGAAAGTGTGCCAGCTTCAAAAACCCGATGTGG 2040  
QY 2041 TGGGTATCAGGAGGCGAGCGGCTTAAGGGCGATGCTCGGGTTTACAAAAGATTAGTGGAA 2100  
Db 2041 TGGGTATCAGGAGGCGAGCGGCTTAAGGGCGATGCTCGGGTTTACAAAAGATTAGTGGAA 2100  
QY 2101 ACATCCTATCTAGCTGCAAGGTTAATTTGATGGAATCTTCCATCTCTGCCGGATGAGTGTG 2160  
Db 2101 ACATCCTATCTAGCTGCAAGGTTAATTTGATGGAATCTTCCATCTCTGCCGGATGAGTGTG 2160  
QY 2161 GGTGTGATACAGCTACATATAAGAGCTGTTATGATGCTTTGATTTTAAAGTTTCATTGGAA 2220  
Db 2161 GGTGTGATACAGCTACATATAAGAGCTGTTATGATGCTTTGATTTTAAAGTTTCATTGGAA 2220  
QY 2221 CTACCAACTTGTTTCTTAAAGAGCTATCTTAAAGCAATATCTCTTTGTTTTTAAACAAA 2280  
Db 2221 CTACCAACTTGTTTCTTAAAGAGCTATCTTAAAGCAATATCTCTTTGTTTTTAAACAAA 2280  
QY 2281 GATATATTTTGTGTATGAAATCTAAATCAAGGCCATCTGTCAATTATGTTACTGTCTTTT 2340  
Db 2281 GATATATTTTGTGTATGAAATCTAAATCAAGGCCATCTGTCAATTATGTTACTGTCTTTT 2340  
QY 2341 TAATCATGTGTTTTGTATATTAATAATTTGTTGACTTCTTAGATTCACTTCCATATGTG 2400  
Db 2341 TAATCATGTGTTTTGTATATTAATAATTTGTTGACTTCTTAGATTCACTTCCATATGTG 2400  
QY 2401 AATGTAAGCTTTAACTATGCTCTTTGTAATGTGTAAATTTCTTCTGAAAATAAACCAT 2460  
Db 2401 AATGTAAGCTTTAACTATGCTCTTTGTAATGTGTAAATTTCTTCTGAAAATAAACCAT 2460  
QY 2461 TTGTGAATAT 2470  
Db 2461 TTGTGAATAT 2470

## RESULT 8

US-10-342-887-1535  
; Sequence 1535, Application US/10342887  
; Publication No. US20040058340A1  
; GENERAL INFORMATION:  
; APPLICANT: Dai, Hongyue  
; APPLICANT: He, Yudong  
; APPLICANT: Linsley, Peter S.  
; APPLICANT: Mao, Mao  
; APPLICANT: Roberts, Christopher J.  
; APPLICANT: Van 't Veer, Laura Johanna  
; APPLICANT: Van de Vijver, Marc J.  
; APPLICANT: Bernarde, Rene  
; TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients  
; FILE REFERENCE: 9301-188-999  
; CURRENT APPLICATION NUMBER: US/10/342,887  
; CURRENT FILING DATE: 2003-01-15  
; PRIOR APPLICATION NUMBER: 60/298,918  
; PRIOR FILING DATE: 2001-06-18  
; PRIOR APPLICATION NUMBER: 60/380,710  
; PRIOR FILING DATE: 2002-05-14  
; PRIOR APPLICATION NUMBER: 10/172,118  
; PRIOR FILING DATE: 2002-06-14  
; NUMBER OF SEQ ID NOS: 2699  
; SEQ ID NO 1535  
; LENGTH: 2470  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-342-887-1535

Query Match 100.0%; Score 2470; DB 8; Length 2470;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2470; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 TTGGCGGGCGGAAGCGGCCACACCCGCGATCGGAAAGATTCTTAGGAACCCCGTACCA 60  
Db 1 TTGGCGGGCGGAAGCGGCCACACCCGCGATCGGAAAGATTCTTAGGAACCCCGTACCA 60



Db 2221 CTACCAACTTGTCTTAAGAGCTATCTTAAGACCAATATCTCTTTGTTTTTAAACAAA 2280  
Qy 2281 GATATATTTTGTGTAGAACTAAATCAAGCCCATCTGTCAATATGTCTCTTTT 2340  
Db 2281 GATATATTTTGTGTAGAACTAAATCAAGCCCATCTGTCAATATGTCTCTTTT 2340  
Qy 2341 TAATCATGTGTTTGTATATTAATTAATTTGTGACCTTCTAGATTCACCTTCATATGTG 2400  
Db 2341 TAATCATGTGTTTGTATATTAATTAATTTGTGACCTTCTAGATTCACCTTCATATGTG 2400  
Qy 2401 AATGTAAGCTCTTAACATATGTCTCTTTGTAATGTGTAATTTCTTCTGAAATAAAACCAT 2460  
Db 2401 AATGTAAGCTCTTAACATATGTCTCTTTGTAATGTGTAATTTCTTCTGAAATAAAACCAT 2460  
Qy 2461 TTGTGAATAT 2470  
Db 2461 TTGTGAATAT 2470

RESULT 9  
US-10-656-598-1  
; Sequence 1, Application US/10656598  
; Publication No. US20040229232A1  
; GENERAL INFORMATION:  
; APPLICANT: DAVID P. DAVIS  
; APPLICANT: FREDERIC J. DESAUVAGE  
; APPLICANT: WILLIAM I. WOOD  
; APPLICANT: ZEMIN ZHANG  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DIAGNOSIS AND  
; TITLE OF INVENTION: TREATMENT OF TUMOR  
; FILE REFERENCE: P1981R1P1-US  
; CURRENT APPLICATION NUMBER: US/10/656,598  
; CURRENT FILING DATE: 2003-09-05  
; PRIOR APPLICATION NUMBER: US 60/410,166  
; PRIOR FILING DATE: 2002-09-11  
; NUMBER OF SEQ ID NOS: 52  
; SEQ ID NO 1  
; LENGTH: 2470  
; TYPE: DNA  
; ORGANISM: Homo sapien  
US-10-656-598-1

Query Match 100.0%; Score 2470; DB 9; Length 2470;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2470; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TTGGCGGGGGAAGCGGCCACCAACCGCGATCGAAAGATCTTAGGAACGCCGTACCA 60  
Db 1 TTGGCGGGGGAAGCGGCCACCAACCGCGATCGAAAGATCTTAGGAACGCCGTACCA 60  
Qy 61 GCCGGTCTCTCAGGACAGCAGCGCCCTGTCTTCTGTGCGGGCGCGCTCAGCCGTGCC 120  
Db 61 GCCGGTCTCTCAGGACAGCAGCGCCCTGTCTTCTGTGCGGGCGCGCTCAGCCGTGCC 120  
Qy 121 TCGGCCCTCAGGTCTTTTCTAATTCAAATAAATCTTGAAGAGACTATGAAGATT 180  
Db 121 TCGGCCCTCAGGTCTTTTCTAATTCCAATAAATCTTGAAGAGACTATGAAGATT 180  
Qy 181 ATGATGAATCTCTCAATAATATTAATTAATGAACTATGGAAGAGACTATGAAGATT 240  
Db 181 ATGATGAATCTCTCAATAATATTAATTAATGAACTATGGAAGAGACTATGAAGATT 240  
Qy 241 AGGTCAAACTTGCCTGCCATATCTTTACTTGGAGAGATGTTAGTATAAAATCATGGATA 300  
Db 241 AGGTCAAACTTGCCTGCCATATCTTTACTTGGAGAGATGTTAGTATAAAATCATGGATA 300  
Qy 301 AAAACACACTAGGAGTGTATTTGCCCGGATCAAAACGGAGATTGAGGCCCTTGAAGACC 360  
Db 301 AAAACACACTAGGAGTGTATTTGCCCGGATCAAAACGGAGATTGAGGCCCTTGAAGACC 360  
Qy 361 TCAGACATCAGCATATATGTCAACTCTACCATGTCTAGACAGCCCAAAATATTTCA 420  
Db 361 TGAGACATCAGCATATATGTCAACTCTACCATGTCTAGACAGCCCAAAATATTTCA 420

Qy 421 TGGTCTTTCAGTACTGCCCTCGAGGAGAGCTGTTGACTATATAAATTTCCAGGATCGCC 480  
Db 421 TGGTCTTTCAGTACTGCCCTCGAGGAGAGCTGTTGACTATATAAATTTCCAGGATCGCC 480  
Qy 481 TGTCAAGAGAGAGACCCGGGTTGCTTCGTCAGATAGTATCTGCTGTATGTGC 540  
Db 481 TGTCAAGAGAGAGACCCGGGTTGCTTCGTCAGATAGTATCTGCTGTATGTGC 540  
Qy 541 ACAGCCAGGCTATGCTCACAGGACCTCAAGCCAGAAAATTTGCTGTTGATGAATATC 600  
Db 541 ACAGCCAGGCTATGCTCACAGGACCTCAAGCCAGAAAATTTGCTGTTGATGAATATC 600  
Qy 601 AFAAATTAAGCTGATTGATCTCTGTGCAAAACCCAAAGGTAACAAGGATTACC 660  
Db 601 AFAAATTAAGCTGATTGATCTCTGTGCAAAACCCAAAGGTAACAAGGATTACC 660  
Qy 661 ATCTACAGACATGCTGTGGGAGCTGTGGCTTATGAGCAGCTGAGTTAATACAGGCAAT 720  
Db 661 ATCTACAGACATGCTGTGGGAGCTGTGGCTTATGAGCAGCTGAGTTAATACAGGCAAT 720  
Qy 721 CATATCTTGGATCAGAGGAGATGTTGGAGCATGGGCATCTGTTATATGTTTATGT 780  
Db 721 CATATCTTGGATCAGAGGAGATGTTGGAGCATGGGCATCTGTTATATGTTTATGT 780  
Qy 781 GTGGAATTTCTACCATTTGATGATGATAATGTAATGCTTTTATCAAGAAGATTATGAGAG 840  
Db 781 GTGGAATTTCTACCATTTGATGATGATAATGTAATGCTTTTATCAAGAAGATTATGAGAG 840  
Qy 841 GAAATATGATGTTCCCAAGTGCTCTCCAGTAGCATCTGCTTCTTCAACAAATGC 900  
Db 841 GAAATATGATGTTCCCAAGTGCTCTCTCCAGTAGCATCTGCTTCTTCAACAAATGC 900  
Qy 901 TGCAGGTGGACCCAAAGAAACCGATTTCTATGAAAAATCTATTGAACCATCTCTGATCA 960  
Db 901 TGCAGGTGGACCCAAAGAAACCGATTTCTATGAAAAATCTATTGAACCATCTCTGATCA 960  
Qy 961 TGCAGGATTACAACTATCTCTGTGAGTGCAAGCAAGAAATCTTTTATTCACCTCGATG 1020  
Db 961 TGCAGGATTACAACTATCTCTGTGAGTGCAAGCAAGAAATCTTTTATTCACCTCGATG 1020  
Qy 1021 ATGATTTGCTAACAACTTTCTGTACATCAGAGAAACAGGCAAAATGAGGATT 1080  
Db 1021 ATGATTTGCTAACAACTTTCTGTACATCAGAGAAACAGGCAAAATGAGGATT 1080  
Qy 1081 TAAATTTCACTGTGGCAGTATGATCACCTCACGGCTACCTATCTCTGCTTCTAGCCAGA 1140  
Db 1081 TAAATTTCACTGTGGCAGTATGATCACCTCACGGCTACCTATCTCTGCTTCTAGCCAGA 1140  
Qy 1141 AGGCTCGGGGAAACCACTGTTTAAAGCTTTTCTTTTCTGTGCAAGCCAGTG 1200  
Db 1141 AGGCTCGGGGAAACCACTGTTTAAAGCTTTTCTTTTCTGTGCAAGCCAGTG 1200  
Qy 1201 CTACCCATTCACAGACATCAAGTCAAAATATTTGAGTCTGGAAGATGAGCCGCAAGTG 1260  
Db 1201 CTACCCATTCACAGACATCAAGTCAAAATATTTGAGTCTGGAAGATGAGCCGCAAGTG 1260  
Qy 1261 AFAAATAATGTTGGCGGATTAATAGACTATGATTTGGTGTGAAGATGATTTATCAACAG 1320  
Db 1261 AFAAATAATGTTGGCGGATTAATAGACTATGATTTGGTGTGAAGATGATTTATCAACAG 1320  
Qy 1321 GTGCTGTACTCCCCGAAACATCACAGTTTACCAAGTACTGGACAGAAATCAAAATGGGTGG 1380  
Db 1321 GTGCTGTACTCCCCGAAACATCACAGTTTACCAAGTACTGGACAGAAATCAAAATGGGTGG 1380  
Qy 1381 AATCTAAATCATTAATCCAGCTTATGCAAGACCTGCAAAATATAAAGAACAAAG 1440  
Db 1381 AATCTAAATCATTAATCCAGCTTATGCAAGACCTGCAAAATATAAAGAACAAAG 1440  
Qy 1441 AAAATGTATATACTCTTAAGTCTGCTGTAAGAAATGAAAGATGATTTATGTTCTCTGAGC 1500  
Db 1441 AAAATGTATATACTCTTAAGTCTGCTGTAAGAAATGAAAGATGATTTATGTTCTCTGAGC 1500

```
QY 1501 CAAAGACTCCAGTTAATAAGAACCCAGCATAGAGAGAAATATCTCATAGCCAAATCGTT 1560
Db 1501 CAAAGACTCCAGTTAATAAGAACCCAGCATAGAGAGAAATATCTCATAGCCAAATCGTT 1560
QY 1561 ACATACACCTCAAAGCTAGAAACCCAGTGCCTGAAAGAACTCCAATTAATAATACCAG 1620
Db 1561 ACATACACCTCAAAGCTAGAAACCCAGTGCCTGAAAGAACTCCAATTAATAATACCAG 1620
QY 1621 TAAATTTCAACAGGAACAGACAAGTTAATGACAGGTGTCTATTAGCCCTGAGAGCGGTGCC 1680
Db 1621 TAAATTTCAACAGGAACAGACAAGTTAATGACAGGTGTCTATTAGCCCTGAGAGCGGTGCC 1680
QY 1681 GCTCAGTGGAAATTTGGATCTCAACCAAGCACATATGGAGGAGACTCCAAAAGAAAGGGAG 1740
Db 1681 GCTCAGTGGAAATTTGGATCTCAACCAAGCACATATGGAGGAGACTCCAAAAGAAAGGGAG 1740
QY 1741 CCAAAGTGTGTGGAGGCTTTGAAAGGGGGTTGGATAAGGTATCACTGTGCTCACCAGGA 1800
Db 1741 CCAAAGTGTGTGGAGGCTTTGAAAGGGGGTTGGATAAGGTATCACTGTGCTCACCAGGA 1800
QY 1801 GCAAAAGGAAGGCTTCTGCCAGAGACGGGCCAGAGACTTAAAGCTTCACTATAATGTGA 1860
Db 1801 GCAAAAGGAAGGCTTCTGCCAGAGACGGGCCAGAGACTTAAAGCTTCACTATAATGTGA 1860
QY 1861 CTACAACTAGATTAGTGAATCCAGATCAACTGTTGAAATGAATTAATGTCTATTCTTCCAA 1920
Db 1861 CTACAACTAGATTAGTGAATCCAGATCAACTGTTGAAATGAATTAATGTCTATTCTTCCAA 1920
QY 1921 AGAAGCATGTGACTTTGTACAAAGGGTTATACATGAAGTGTCAACACAGTCAAGTT 1980
Db 1921 AGAAGCATGTGACTTTGTACAAAGGGTTATACATGAAGTGTCAACACAGTCAAGTT 1980
QY 1981 TTGGGAAAGTGACAATGCAATTTGAATTAGAAGTGTGCCAGCTTCAAAAACCCGATGTGG 2040
Db 1981 TTGGGAAAGTGACAATGCAATTTGAATTAGAAGTGTGCCAGCTTCAAAAACCCGATGTGG 2040
QY 2041 TGGGTATCAGGAGCGAGCGCTTAAGGGCGATCGCTGGGTTTACAAAAGATTAGTGGAG 2100
Db 2041 TGGGTATCAGGAGCGAGCGCTTAAGGGCGATCGCTGGGTTTACAAAAGATTAGTGGAG 2100
QY 2101 ACATCTATCTAGTGCAGAGTATAATTGATGGAATCTTCCATCTGCCGGATGAGTGTG 2160
Db 2101 ACATCTATCTAGTGCAGAGTATAATTGATGGAATCTTCCATCTGCCGGATGAGTGTG 2160
QY 2161 GGTGTGATACAGCTCATATAAGACTGTATGATCGCTTTGATTTAAAGTTTCATTGGAA 2220
Db 2161 GGTGTGATACAGCTCATATAAGACTGTATGATCGCTTTGATTTAAAGTTTCATTGGAA 2220
QY 2221 CTACCAACTTGTCTTAAAGAGCTATCTTAAGACCAATATCTCTTTGTTTTTAAACAAA 2280
Db 2221 CTACCAACTTGTCTTAAAGAGCTATCTTAAGACCAATATCTCTTTGTTTTTAAACAAA 2280
QY 2281 GATATTATTTTGTGATGAATCTTAAATCAAGCCCATCTGTCATTAATGTTACTGCTTTTT 2340
Db 2281 GATATTATTTTGTGATGAATCTTAAATCAAGCCCATCTGTCATTAATGTTACTGCTTTTT 2340
QY 2341 TAATCATGTGGTTTGTGATATAATATCTTGAATCTTCTTAGATTCACATTCATATGTG 2400
Db 2341 TAATCATGTGGTTTGTGATATAATATCTTGAATCTTCTTAGATTCACATTCATATGTG 2400
QY 2401 AATGTAAGCTCTTAACTATGCTCTTTTGAATGTGAATTTCTTTCTGAAATAAAACCAT 2460
Db 2401 AATGTAAGCTCTTAACTATGCTCTTTTGAATGTGAATTTCTTTCTGAAATAAAACCAT 2460
QY 2461 TTGTGAATAT 2470
Db 2461 TTGTGAATAT 2470
```

RESULT 10

US-10-723-860-2443

; Sequence 2443, Application US/10723860

; Publication No. US20040253606A1

```
; GENERAL INFORMATION:
; APPLICANT: Aziz, Natasha
; APPLICANT: Ginsburg, Wendy M.
; APPLICANT: Zlotnik, Albert
; TITLE OF INVENTION: Methods of Diagnosis of Soft Tissue Sarcoma, Compositions &
; FILE REFERENCE: 05882.0193.NPUS01
; CURRENT FILING DATE: 2003-11-26
; PRIOR FILING DATE: 2002-11-26
; NUMBER OF SEQ ID NOS: 8393
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 2443
; LENGTH: 2470
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-723-860-2443
```

```
Query Match 100.0%; Score 2470; DB 9; Length 2470;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2470; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTGGCGGGCGGAAGCGGCCACAAACCCGGCGATCGAAAAGATTCTTAGGAACGGCGTACCA 60
Db 1 TTGGCGGGCGGAAGCGGCCACAAACCCGGCGATCGAAAAGATTCTTAGGAACGGCGTACCA 60
QY 61 GCGCGTCTCTCAGGACAGCAGCGCCCTGTCTCTGTGGGGCGCGCTCAGCGCGTGC 120
Db 61 GCGCGTCTCTCAGGACAGCAGCGCCCTGTCTCTGTGGGGCGCGCTCAGCGCGTGC 120
QY 121 TCGCGCCCTCAGGTCTCTTTCTTAATTCAAATAAATCTTGAAGAGACTATCAAGAGATT 180
Db 121 TCGCGCCCTCAGGTCTCTTTCTTAATTCAAATAAATCTTGAAGAGACTATCAAGAGATT 180
QY 181 ATGATGAATCTTCAATAATATTATGAATTACATGAATACTATTGGACAGGTGGCTTGCAA 240
Db 181 ATGATGAATCTTCAATAATATTATGAATTACATGAATACTATTGGACAGGTGGCTTGCAA 240
QY 241 AGGTCAAACTTGCTGCCATATCTTACTGGAGAGATGGTAGCTATAAAATCATGGATA 300
Db 241 AGGTCAAACTTGCTGCCATATCTTACTGGAGAGATGGTAGCTATAAAATCATGGATA 300
QY 301 AAAACACACTAGGAGGTGATTTGCCCGGATCAAAAACGGAGATTGAGGCTTCAAGAACC 360
Db 301 AAAACACACTAGGAGGTGATTTGCCCGGATCAAAAACGGAGATTGAGGCTTCAAGAACC 360
QY 361 TGAGACATCAGCATATATGTCAACTCTACCATGTCTAGACAGAGCCACAAAATATTCA 420
Db 361 TGAGACATCAGCATATATGTCAACTCTACCATGTCTAGACAGAGCCACAAAATATTCA 420
QY 421 TGCTTCTTGAGTACTGCCCTGGAGGAGAGCTGTTTGACTATATAATTTCCAGGATCGCC 480
Db 421 TGCTTCTTGAGTACTGCCCTGGAGGAGAGCTGTTTGACTATATAATTTCCAGGATCGCC 480
QY 481 TGTCAAGAAGGAGACCCGGGTTGTCTTCGTGAGATAGTATCTGCTGTTGCTTATGTGC 540
Db 481 TGTCAAGAAGGAGACCCGGGTTGTCTTCGTGAGATAGTATCTGCTGTTGCTTATGTGC 540
QY 541 ACAGCCAGGCTATGTCTCAGGGACCTCAAGCGAGAAAATTTGCTGTTTGTATGATATC 600
Db 541 ACAGCCAGGCTATGTCTCAGGGACCTCAAGCGAGAAAATTTGCTGTTTGTATGATATC 600
QY 601 ATAAATTTAAAGCTGATTGACTTTGGTCTCTGTGCAAAAACCAAGGGTAAACAAGATTACC 660
Db 601 ATAAATTTAAAGCTGATTGACTTTGGTCTCTGTGCAAAAACCAAGGGTAAACAAGATTACC 660
QY 661 ATCTACAGACATGCTGTGGAGTCTGGCTTATGAGCAGCACTGTTTAAATCAAGGCAAT 720
Db 661 ATCTACAGACATGCTGTGGAGTCTGGCTTATGAGCAGCACTGTTTAAATCAAGGCAAT 720
QY 721 CATATCTTGGATCAGAGGCGAGATGTTTGGAGCATGGCATCTGTTATATGTTCTTATGT 780
Db 721 CATATCTTGGATCAGAGGCGAGATGTTTGGAGCATGGCATCTGTTATATGTTCTTATGT 780
```



Db 721 CATATCTTGGATCAGAGGACAGATGTTTGGAGCATGGGCATACCTGTTATATATGTTCTTATGT 780  
Qy 781 GTGGATTTCTACATTTGATGATGAATAATTAATGCTTTTATACAAAGAGATTATGAGAG 840  
Db 781 GTGGATTTCTACATTTGATGATGAATAATTAATGCTTTTATACAAAGAGATTATGAGAG 840  
Qy 841 GAAAAATATGATGTTCCCAAGTGGCTCTCTCCAGTAGTAGCTTCTTCTTCAACAATGC 900  
Db 841 GAAAAATATGATGTTCCCAAGTGGCTCTCTCCAGTAGTAGCTTCTTCTTCAACAATGC 900  
Qy 901 TGCAGGTGGACCCCAAGAAACGGATTCTATGAAAAATCTATGAAACCATCCCTCGATCA 960  
Db 901 TGCAGGTGGACCCCAAGAAACGGATTCTATGAAAAATCTATGAAACCATCCCTCGATCA 960  
Qy 961 TGCAGGATTACAACTATCTCTGTTGAGTGGCAAGCAAGAAATCTTTTATTTACCTCGATG 1020  
Db 961 TGCAGGATTACAACTATCTCTGTTGAGTGGCAAGCAAGAAATCTTTTATTTACCTCGATG 1020  
Qy 1021 ATGATTGCGTAAACAGAACTTTCTGTACATCACAGAAACACAGGCAAAACAATGGAGATT 1080  
Db 1021 ATGATTGCGTAAACAGAACTTTCTGTACATCACAGAAACACAGGCAAAACAATGGAGATT 1080  
Qy 1081 TAAATTTCACTGTGGCAGTATGATCACTCACGGCTACCTATCTCTGCTTCTAGCCAAAG 1140  
Db 1081 TAAATTTCACTGTGGCAGTATGATCACTCACGGCTACCTATCTCTGCTTCTAGCCAAAG 1140  
Qy 1141 AGCTCGGGGAAACCAAGTTCGTTTAAAGGCTTTCTTTCTCTCTGTGACCAAGCCAGTG 1200  
Db 1141 AGCTCGGGGAAACCAAGTTCGTTTAAAGGCTTTCTTTCTCTCTGTGACCAAGCCAGTG 1200  
Qy 1201 CTACCCCATTCACAGACATCAAGTCAAAATTAATGGAGTCTGGAAGATGACCGCAAGTG 1260  
Db 1201 CTACCCCATTCACAGACATCAAGTCAAAATTAATGGAGTCTGGAAGATGACCGCAAGTG 1260  
Qy 1261 ATAAAAATTTATGTGGCGGATTAATAGACTATGATTGGTGTGAAGATGATTATCAACAG 1320  
Db 1261 ATAAAAATTTATGTGGCGGATTAATAGACTATGATTGGTGTGAAGATGATTATCAACAG 1320  
Qy 1321 GTGCTGCTACTCCCGAATCATCAGTTTACCAAGTACTGGACAGAAATCAAAATGGGGTGG 1380  
Db 1321 GTGCTGCTACTCCCGAATCATCAGTTTACCAAGTACTGGACAGAAATCAAAATGGGGTGG 1380  
Qy 1381 AATCTAAATCATTAATCTCAGCCTTATGACAGACACCTGCAAAATTAATTAAGAACAAAG 1440  
Db 1381 AATCTAAATCATTAATCTCAGCCTTATGACAGACACCTGCAAAATTAATTAAGAACAAAG 1440  
Qy 1441 AAAATGTATATCTCTTAAGTCTGCTGTAAAGAAATGAAGAGTACTTTATGTTTCTCTGAGC 1500  
Db 1441 AAAATGTATATCTCTTAAGTCTGCTGTAAAGAAATGAAGAGTACTTTATGTTTCTCTGAGC 1500  
Qy 1501 CAAAGACTCCAGTTAATAAGAACCCAGCATATAAGAGAGAAATCTCACTACGCCAAATCGTT 1560  
Db 1501 CAAAGACTCCAGTTAATAAGAACCCAGCATATAAGAGAGAAATCTCACTACGCCAAATCGTT 1560  
Qy 1561 ACATACACCTCAAAAGCTAGAAACCAAGTCTGAAAGAAATCTCAATTAATAAATACCAAG 1620  
Db 1561 ACATACACCTCAAAAGCTAGAAACCAAGTCTGAAAGAAATCTCAATTAATAAATACCAAG 1620  
Qy 1621 TAAATTCACAGGAAACAGACAGTTAATCAGAGGTGCTATTAGCCCTGAGAGCGGTGCC 1680  
Db 1621 TAAATTCACAGGAAACAGACAGTTAATCAGAGGTGCTATTAGCCCTGAGAGCGGTGCC 1680  
Qy 1681 GCTCAGTGAATTTGGAATCTCAACCAAGCACATATGGAGAGACTCCAAAAAGAAAGGGAG 1740  
Db 1681 GCTCAGTGAATTTGGAATCTCAACCAAGCACATATGGAGAGACTCCAAAAAGAAAGGGAG 1740  
Qy 1741 CCAAAGTGTGGAGCCTTTGAAAGGGGTTGGATAAGTTATCACTGTGCTCACCAGGA 1800  
Db 1741 CCAAAGTGTGGAGCCTTTGAAAGGGGTTGGATAAGTTATCACTGTGCTCACCAGGA 1800  
Qy 1801 GCAAAGGAAGGGTCTGCGCAGAGACGGGCCAGAGACTAAAGCTTCACTATATATGGA 1860  
Db 1801 GCAAAGGAAGGGTCTGCGCAGAGACGGGCCAGAGACTAAAGCTTCACTATATATGGA 1860

Qy 1861 CTACAACTAGATTAGTGAATCCAGATCAACTGTGTGAATGAAATTAATGTCTTATTTCCAA 1920  
Db 1861 CTACAACTAGATTAGTGAATCCAGATCAACTGTGTGAATGAAATTAATGTCTTATTTCCAA 1920  
Qy 1921 AGAAGCATCTTGCATTTGTACAAAAGGGTTATACACTGAAGTGTCAACACAGTCAGATT 1980  
Db 1921 AGAAGCATCTTGCATTTGTACAAAAGGGTTATACACTGAAGTGTCAACACAGTCAGATT 1980  
Qy 1981 TTGGGAAAGTGACAAATGCAATTTGAATTAGAAAGTGTGCCAGCTTCAAAAACCCGATGTGG 2040  
Db 1981 TTGGGAAAGTGACAAATGCAATTTGAATTAGAAAGTGTGCCAGCTTCAAAAACCCGATGTGG 2040  
Qy 2041 TGGGTATCAGAGGCGAGCGGCTTAAAGGCGATGCTCGGGTTTACAAAAGATTAGTGAAG 2100  
Db 2041 TGGGTATCAGAGGCGAGCGGCTTAAAGGCGATGCTCGGGTTTACAAAAGATTAGTGAAG 2100  
Qy 2101 ACATCTCTATCTAGCTGCAAGGTATTAATTCATGGATTCTTCCATCTGCGCGGATGAGTGG 2160  
Db 2101 ACATCTCTATCTAGCTGCAAGGTATTAATTCATGGATTCTTCCATCTGCGCGGATGAGTGG 2160  
Qy 2161 GGTGTGATACAGCTACATAAAGACTGTTATGATGCTTTGATTTTAAAGTTCATTGGAA 2220  
Db 2161 GGTGTGATACAGCTACATAAAGACTGTTATGATGCTTTGATTTTAAAGTTCATTGGAA 2220  
Qy 2221 CTACCAACTGTTTCTTAAAGAGCTATCTTAAAGCAATATCTCTTTGTTTTTAAACAAA 2280  
Db 2221 CTACCAACTGTTTCTTAAAGAGCTATCTTAAAGCAATATCTCTTTGTTTTTAAACAAA 2280  
Qy 2281 GATATTTATTTGTGTATGATCTAAATCAAGCCCATCTGCTCATATGTTACTGTCTTTTT 2340  
Db 2281 GATATTTATTTGTGTATGATCTAAATCAAGCCCATCTGCTCATATGTTACTGTCTTTTT 2340  
Qy 2341 TAAATCATGTTTGTATATTAATTAATGTTGACTTTCTTAGATTCACTTCCATATGTG 2400  
Db 2341 TAAATCATGTTTGTATATTAATTAATGTTGACTTTCTTAGATTCACTTCCATATGTG 2400  
Qy 2401 AATGAAGCTCTTAACTATGCTCTTTGTAATGTTGTAATTTCTTTCTGAAATAAACCAT 2460  
Db 2401 AATGAAGCTCTTAACTATGCTCTTTGTAATGTTGTAATTTCTTTCTGAAATAAACCAT 2460  
Qy 2461 TTGTGAATAT 2470  
Db 2461 TTGTGAATAT 2470

RESULT 11  
US-10-848-755A-169  
; Sequence 169, Application US/10848755A  
; Publication No. US20050054826A1  
; GENERAL INFORMATION:  
; APPLICANT: Mao, Mao  
; TITLE OF INVENTION: HUMAN DIAPHANOUS-3 GENE AND METHODS OF USE THEREFOR  
; FILE REFERENCE: 9301-196-999  
; CURRENT APPLICATION NUMBER: US/10/848,755A  
; CURRENT FILING DATE: 2004-05-18  
; PRIOR APPLICATION NUMBER: 60/471,842  
; PRIOR FILING DATE: 2003-05-11  
; NUMBER OF SEQ ID NOS: 275  
; SOFTWARE: PatentIn version 3.2 CAM: 301891-999188  
; SEQ ID NO 169  
; LENGTH: 2470  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-848-755A-169

Query Match 100.0%; Score 2470; DB 10; Length 2470;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2470; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 TTGGCGGGCGGAAGCGGCCAACACCCGGCGATCGAAAGATCTTTAGGAACGCCGTACCA 60  
Db 1 TTGGCGGGCGGAAGCGGCCAACACCCGGCGATCGAAAGATCTTTAGGAACGCCGTACCA 60



QY 61 GCGGCTCTCTCAGGACAGACGCCCTCTGCTTCTGTCGGCGCGCTCAGCGTGCC 120  
Db 61 GCGGCTCTCTCAGGACAGACGCCCTCTGCTTCTGTCGGCGCGCTCAGCGTGCC 120  
QY 121 TCGGCCCTCAGGTTCTTTTCTTAATTCAAATFAAACTTGCAGAGGACTATGAAGATT 180  
Db 121 TCGGCCCTCAGGTTCTTTTCTTAATTCAAATFAAACTTGCAGAGGACTATGAAGATT 180  
QY 181 ATGATGAACCTTCTCAAATATATGAATATACATGAATATATGGGACAGGTGGCTTGC 240  
Db 181 ATGATGAACCTTCTCAAATATATGAATATACATGAATATATGGGACAGGTGGCTTGC 240  
QY 241 AGGTCAAACTTGCCTGCCATATCTTACTTGGAGAGATGTTAGCTATAAAAAATCATGGATA 300  
Db 241 AGGTCAAACTTGCCTGCCATATCTTACTTGGAGAGATGTTAGCTATAAAAAATCATGGATA 300  
QY 301 AAAACACACTAGGAGTGTATTTGCCCGGATCAAAAACGAGATGAGGCTTTGAAGAACC 360  
Db 301 AAAACACACTAGGAGTGTATTTGCCCGGATCAAAAACGAGATGAGGCTTTGAAGAACC 360  
QY 361 TGAGACATCAGCATATATGTCAACTCTACATGTGCTAGAGACAGCAACAAAATATTC 420  
Db 361 TGAGACATCAGCATATATGTCAACTCTACATGTGCTAGAGACAGCAACAAAATATTC 420  
QY 421 TGGTCTTGAGTACTGCCCTGGAGGAGAGCTGTTGACTATATAATTTCCAGGATCGCC 480  
Db 421 TGGTCTTGAGTACTGCCCTGGAGGAGAGCTGTTGACTATATAATTTCCAGGATCGCC 480  
QY 481 TGTCAAGAGAGAGACCCGGTGTCTTCCGTGAGATAGTATCTGCTGTTGCTTATGTGC 540  
Db 481 TGTCAAGAGAGAGACCCGGTGTCTTCCGTGAGATAGTATCTGCTGTTGCTTATGTGC 540  
QY 541 ACAGCCAGGGCTATGTCTCAGGGAACCTCAAGCCAGAAAATTTGCTGTTGATGAATATC 600  
Db 541 ACAGCCAGGGCTATGTCTCAGGGAACCTCAAGCCAGAAAATTTGCTGTTGATGAATATC 600  
QY 601 ATAAATTAAGCTGATGATTTGCTGCTGTGCAAAACCCAGGGTAAACAGGATTACC 660  
Db 601 ATAAATTAAGCTGATGATTTGCTGCTGTGCAAAACCCAGGGTAAACAGGATTACC 660  
QY 661 ATCTACAGACATGCTGTGGAGTCTGGCTTATGACGACCTGAGTTAATACAGGCAAT 720  
Db 661 ATCTACAGACATGCTGTGGAGTCTGGCTTATGACGACCTGAGTTAATACAGGCAAT 720  
QY 721 CATATCTTGATCAGAGCGCATGTTTGGAGCATGGGCATCTGTTATATGTTCTTATGT 780  
Db 721 CATATCTTGATCAGAGCGCATGTTTGGAGCATGGGCATCTGTTATATGTTCTTATGT 780  
QY 781 GTGGATTTCTACCATTTGATGATAATGTAATGCTTTATACAAAGAGATTATCAGAG 840  
Db 781 GTGGATTTCTACCATTTGATGATAATGTAATGCTTTATACAAAGAGATTATCAGAG 840  
QY 841 GAAATATATGATGTTCCCAAGTGGCTCTCTCCAGTAGCATTTGCTTCTTCAACAAATGC 900  
Db 841 GAAATATATGATGTTCCCAAGTGGCTCTCTCCAGTAGCATTTGCTTCTTCAACAAATGC 900  
QY 901 TGCAGGTGGACCCAAAGAAACGGATTTCTATGAAAAATCTATTTGAACCATCCCTGGATCA 960  
Db 901 TGCAGGTGGACCCAAAGAAACGGATTTCTATGAAAAATCTATTTGAACCATCCCTGGATCA 960  
QY 961 TSCAAGATTAACAATCTCCTGTTGAGTGGCAAGCAAGATCCTTTTATTCACCTCGATG 1020  
Db 961 TSCAAGATTAACAATCTCCTGTTGAGTGGCAAGCAAGATCCTTTTATTCACCTCGATG 1020  
QY 1021 ATGATTTGCGTAAACAGAACTTTCTGTACATCACAGAAAACACAGGCAAAACAAATGGAGATT 1080  
Db 1021 ATGATTTGCGTAAACAGAACTTTCTGTACATCACAGAAAACACAGGCAAAACAAATGGAGATT 1080  
QY 1081 TAAATTCATGTGGGAGTATGATCACTCAAGGCTACCTATCTTCTGCTTCTAGCCAGA 1140  
Db 1081 TAAATTCATGTGGGAGTATGATCACTCAAGGCTACCTATCTTCTGCTTCTAGCCAGA 1140

QY 1141 AGGCTCGGGGAAAAACAGTTCGTTTAAAGCTTTCTTTCTCTGTGAGCAAGCCAGTG 1200  
Db 1141 AGGCTCGGGGAAAAACAGTTCGTTTAAAGCTTTCTTTCTCTGTGAGCAAGCCAGTG 1200  
QY 1201 CTACCCCATTTACAGACATCAAGTCAAAATATTTGAGTCTGGAAGATGTGACCGCAAGTG 1260  
Db 1201 CTACCCCATTTACAGACATCAAGTCAAAATATTTGAGTCTGGAAGATGTGACCGCAAGTG 1260  
QY 1261 ATAAAAAATTTATGTGCGGGATTTAATAGACTATGATTTGGTGTGAAGATGATTTATCAACAG 1320  
Db 1261 ATAAAAAATTTATGTGCGGGATTTAATAGACTATGATTTGGTGTGAAGATGATTTATCAACAG 1320  
QY 1321 GTGCTGTACTCCCGAACATCACAGTTTACCAAGTACTGGAAGATGATTTATGGGTGG 1380  
Db 1321 GTGCTGTACTCCCGAACATCACAGTTTACCAAGTACTGGAAGATGATTTATGGGTGG 1380  
QY 1381 AATCTAAATCATTAATCTCAGGCTTATGCAGAACACCTGCAAAATAAATTTAAAGAACAAAG 1440  
Db 1381 AATCTAAATCATTAATCTCAGGCTTATGCAGAACACCTGCAAAATAAATTTAAAGAACAAAG 1440  
QY 1441 AAAATGTATATACTCTTAAGTCTGCTGTAAGAAATGAAGAGTACTTTATGTTTCTCTGAGC 1500  
Db 1441 AAAATGTATATACTCTTAAGTCTGCTGTAAGAAATGAAGAGTACTTTATGTTTCTCTGAGC 1500  
QY 1501 CAAAGACTCCAGTTAATAAGAACCCAGCATAAAGAGAAATACTCACTACGCCAAATCGTT 1560  
Db 1501 CAAAGACTCCAGTTAATAAGAACCCAGCATAAAGAGAAATACTCACTACGCCAAATCGTT 1560  
QY 1561 ACATACACCTTCAAAAGCTAGAAAACCAAGTCTGCTGAAAGAAATCTCAATTTAAATACCAAG 1620  
Db 1561 ACATACACCTTCAAAAGCTAGAAAACCAAGTCTGCTGAAAGAAATCTCAATTTAAATACCAAG 1620  
QY 1621 TAAATTTCAACAGAAACAGACAAGTTAATGACAGTGTCTATTAGCCTGAGAGCGGTGCC 1680  
Db 1621 TAAATTTCAACAGAAACAGACAAGTTAATGACAGTGTCTATTAGCCTGAGAGCGGTGCC 1680  
QY 1681 GCTCAGTGAATTTGGATCTCAAACCAAGCATAATATGAGAGAGACTCCAAAAGAAAGGGAG 1740  
Db 1681 GCTCAGTGAATTTGGATCTCAAACCAAGCATAATATGAGAGAGACTCCAAAAGAAAGGGAG 1740  
QY 1741 CCAAAGTGTGGAGCCTTGAAGGGGTTCGTAAGGTTATCACTGTCTCACCAGCA 1800  
Db 1741 CCAAAGTGTGGAGCCTTGAAGGGGTTCGTAAGGTTATCACTGTCTCACCAGCA 1800  
QY 1801 GCAAAGGAAGGTTCTGCCAGAGACGGGCCCAGAAAGACTTAAAGCTTCACTATAATGTGA 1860  
Db 1801 GCAAAGGAAGGTTCTGCCAGAGACGGGCCCAGAAAGACTTAAAGCTTCACTATAATGTGA 1860  
QY 1861 CTACAACTAGATTAGTGAATCCAGATCAACTGTTGAATGAAATAATGTCTATTTCTTCAA 1920  
Db 1861 CTACAACTAGATTAGTGAATCCAGATCAACTGTTGAATGAAATAATGTCTATTTCTTCAA 1920  
QY 1921 AGAAGCATGTTGACTTTGTACAAAAGGTTATACACTGAGGTGTCAAAACACAGTCAAGATT 1980  
Db 1921 AGAAGCATGTTGACTTTGTACAAAAGGTTATACACTGAGGTGTCAAAACACAGTCAAGATT 1980  
QY 1981 TTGGGAAAGTGACAATGCAATTTGAATTAAGAAAGTGTGCCAGCTTCAAAAACCCGATGTGG 2040  
Db 1981 TTGGGAAAGTGACAATGCAATTTGAATTAAGAAAGTGTGCCAGCTTCAAAAACCCGATGTGG 2040  
QY 2041 TGGGTATCAGGAGGCGGCTTAAAGGCGATGCTGGGTTTACAAAAGATTAGTGGAG 2100  
Db 2041 TGGGTATCAGGAGGCGGCTTAAAGGCGATGCTGGGTTTACAAAAGATTAGTGGAG 2100  
QY 2101 ACATCCTATCTAGTCAAGGTTAATTTGATGAGTCTTCCATCTCGCGGATGAGTGTG 2160  
Db 2101 ACATCCTATCTAGTCAAGGTTAATTTGATGAGTCTTCCATCTCGCGGATGAGTGTG 2160  
QY 2161 GGTGTGATACAGCTACATAAAGACTGTTATGATCGCTTTTGAATTTTAAAGTTCATTGGAA 2220  
Db 2161 GGTGTGATACAGCTACATAAAGACTGTTATGATCGCTTTTGAATTTTAAAGTTCATTGGAA 2220  
QY 2221 CTACCAACTTGTCTTAAAGAGCTATCTTTAAGACCAATATCTCTTGTGTTTTTAAACAAA 2280

Db 2221 CTACCACTTGTCTTAAAGAGCTATCTTAAGCAATATCTCTTGTGTTTTTAAACAAA 2280  
Qy 2281 GATATATTATTTGTATGAATCTAAATCAAGCCCATCTGTCAATATGTTACTGCTTTTT 2340  
Db 2281 GATATATTATTTGTATGAATCTAAATCAAGCCCATCTGTCAATATGTTACTGCTTTTT 2340  
Qy 2341 TAATCATGTGGTTTTGTATTAATTAATTTGTCATCTTTTGTAGATCACTTCCATATGTG 2400  
Db 2341 TAATCATGTGGTTTTGTATTAATTAATTTGTCATCTTTTGTAGATCACTTCCATATGTG 2400  
Qy 2401 AATGTAAGCTCTTAACATATGCTCTTTGTAATGTAATTTCTTCTGAAATAAACCAT 2460  
Db 2401 AATGTAAGCTCTTAACATATGCTCTTTGTAATGTAATTTCTTCTGAAATAAACCAT 2460  
Qy 2461 TTGTGAATAT 2470  
Db 2461 TTGTGAATAT 2470

## RESULT 12

US-10-843-641A-4148  
; Sequence 4148, Application US/10843641A  
; Publication No. US20050064454A1  
; GENERAL INFORMATION:  
; APPLICANT: Avalon Pharmaceuticals, Inc.  
; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using  
; FILE REFERENCE: Signature Gene Sets  
; CURRENT APPLICATION NUMBER: US/10/843,641A  
; CURRENT FILING DATE: 2004-05-12  
; PRIOR APPLICATION NUMBER: US/09/873,367  
; PRIOR FILING DATE: 2001-06-05  
; PRIOR APPLICATION NUMBER: US/09/954,531  
; PRIOR FILING DATE: 2001-09-18  
; PRIOR APPLICATION NUMBER: US/09/954,456  
; PRIOR FILING DATE: 2001-09-25  
; PRIOR APPLICATION NUMBER: US/09/962,436  
; PRIOR FILING DATE: 2001-09-25  
; PRIOR APPLICATION NUMBER: US/09/962,832  
; PRIOR FILING DATE: 2001-09-25  
; PRIOR APPLICATION NUMBER: US/09/964,824  
; PRIOR FILING DATE: 2001-09-27  
; PRIOR APPLICATION NUMBER: US/09/967,768  
; PRIOR FILING DATE: 2001-09-28  
; PRIOR APPLICATION NUMBER: US/09/968,007  
; PRIOR FILING DATE: 2001-10-02  
; PRIOR APPLICATION NUMBER: US/09/969,347  
; PRIOR FILING DATE: 2001-10-02  
; PRIOR APPLICATION NUMBER: US/09/969,708  
; PRIOR FILING DATE: 2001-10-03

; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 8447  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 4148  
; LENGTH: 2470  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-843-641A-4148

Query Match 100.0%; Score 2470; DB 10; Length 2470;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2470; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TTGGGGGGGGAAGCGGCCACACCGCGGATCGAAAGATCTTAGGAACCGCGTACCA 60  
Db 1 TTGGGGGGGGAAGCGGCCACACCGCGGATCGAAAGATCTTAGGAACCGCGTACCA 60  
Qy 61 GCCGGGTCTCTCAGGACAGCAGCGCCCTGCTCTTCTGTGCGGCGCGCTCAGCGGTGCC 120  
Db 61 GCCGGGTCTCTCAGGACAGCAGCGCCCTGCTCTTCTGTGCGGCGCGCTCAGCGGTGCC 120  
Qy 121 TCCGCCCTCAGGTTCTTTTCTAAATTCAAATAAACTTGCAAGAGGACTATGAAGATT 180

Db 121 TCCGCCCTCAGGTTCTTTTCTAAATTCAAATAAACTTGCAAGAGGACTATGAAGATT 180  
Qy 181 ATGATGAACCTTCTCAAAATATTATGAATACATGAATACTATGGGACAGGTGGCTTTGCCAA 240  
Db 181 ATGATGAACCTTCTCAAAATATTATGAATACATGAATACTATGGGACAGGTGGCTTTGCCAA 240  
Qy 241 AGGTCAAACTTGCCTGCCATATCTTTACTGGAGAGATGCTATGAATACTATGAATACTATGA 300  
Db 241 AGGTCAAACTTGCCTGCCATATCTTTACTGGAGAGATGCTATGAATACTATGAATACTATGA 300  
Qy 301 AAAACACACTAGGAGTGAATTTGCCCGGATCAAAACGAGATTTAGGCTTTGAAGAAC 360  
Db 301 AAAACACACTAGGAGTGAATTTGCCCGGATCAAAACGAGATTTAGGCTTTGAAGAAC 360  
Qy 361 TGAGACATCAGCATATATGTCACACTCTACCATGCTGTAGAGACAGCAACAAATATTCA 420  
Db 361 TGAGACATCAGCATATATGTCACACTCTACCATGCTGTAGAGACAGCAACAAATATTCA 420  
Qy 421 TGGTCTTCTGAGTACTGCCCTGGAGGAGCTGTTTGACTATATAATTTCCAGGATCGCC 480  
Db 421 TGGTCTTCTGAGTACTGCCCTGGAGGAGCTGTTTGACTATATAATTTCCAGGATCGCC 480  
Qy 481 TGTCAAGAGAGAGACCGGGTGTCTTCCGTGATAGTATCTGCTGTTGCTTATGTC 540  
Db 481 TGTCAAGAGAGAGACCGGGTGTCTTCCGTGATAGTATCTGCTGTTGCTTATGTC 540  
Qy 541 ACAGCCAGGGCTATGCTCACAGGGACCTCAAGCCAGAAAATTTGCTGTTGATGAATATC 600  
Db 541 ACAGCCAGGGCTATGCTCACAGGGACCTCAAGCCAGAAAATTTGCTGTTGATGAATATC 600  
Qy 601 ATAAATTAAGCTGATGTAATTTGGTCTCTGTGCAAAACCCAAAGGTAAACAGGATTACC 660  
Db 601 ATAAATTAAGCTGATGTAATTTGGTCTCTGTGCAAAACCCAAAGGTAAACAGGATTACC 660  
Qy 661 ATCTACAGACATGCTGTGGAGTCTGGCTTATGACAGCACCCTGAGTTAATACAGGCAAT 720  
Db 661 ATCTACAGACATGCTGTGGAGTCTGGCTTATGACAGCACCCTGAGTTAATACAGGCAAT 720  
Qy 721 CATATCTTGGATCAGAGCAGATGTTGGAGCATGGGCATCTGTTATATGTTCTTATGT 780  
Db 721 CATATCTTGGATCAGAGCAGATGTTGGAGCATGGGCATCTGTTATATGTTCTTATGT 780  
Qy 781 GTGGATTTCTACCAATTTGATGATGATAATGTAATGGCTTTATACAAAGAGATTATGAGAG 840  
Db 781 GTGGATTTCTACCAATTTGATGATGATAATGTAATGGCTTTATACAAAGAGATTATGAGAG 840  
Qy 841 GAAATATGATGTTCCCAAGTGGCTCTCTCCAGTAGCATTTCTGCTTCTTCAACAAATGC 900  
Db 841 GAAATATGATGTTCCCAAGTGGCTCTCTCCAGTAGCATTTCTGCTTCTTCAACAAATGC 900  
Qy 901 TGCAGGTGACCCCAAGAAAACGGATTTCTATGAAAATCTATTTGAACCATCTCTGGATCA 960  
Db 901 TGCAGGTGACCCCAAGAAAACGGATTTCTATGAAAATCTATTTGAACCATCTCTGGATCA 960  
Qy 961 TGCAAGATTACAACTATCTCTGTTGAGTGGCAAGCAAGATCTCTTTTATTCACCTCGATG 1020  
Db 961 TGCAAGATTACAACTATCTCTGTTGAGTGGCAAGCAAGATCTCTTTTATTCACCTCGATG 1020  
Qy 1021 ATGATTGGGTAAACAGAACTTTCTGTACATCAGAAAACAGGCAACAAATGGAGGATT 1080  
Db 1021 ATGATTGGGTAAACAGAACTTTCTGTACATCAGAAAACAGGCAACAAATGGAGGATT 1080  
Qy 1081 TAAATTTCACTGTGGCAGTATGATACCTCAGGCTACCTATCTCTGCTTCTGCTTCTAGCAAGA 1140  
Db 1081 TAAATTTCACTGTGGCAGTATGATACCTCAGGCTACCTATCTCTGCTTCTGCTTCTAGCAAGA 1140  
Qy 1141 AGGCTCGGGGAAAACAGTTCGTTTAAAGCTTTCTTTCTCTGTGTGCAAGCCAGTG 1200  
Db 1141 AGGCTCGGGGAAAACAGTTCGTTTAAAGCTTTCTTTCTCTGTGTGCAAGCCAGTG 1200  
Qy 1201 CTACCCCATTCACAGACATCAAGTCAATTAATTTGAGTCTGGAAGATGTGACCGCAAGTG 1260

Db 1201 CTACCCATTACAGACATCAAGTCAAAATAATTGGAGTCTGGAAGATGTGACCGCAAGTG 1260

Qy 1261 ATAAAAATTATGTGGCGGATTAATAGACTATGATTGGTGTGAAGATGATTATCAACAG 1320

Db 1261 ATAAAAATTATGTGGCGGATTAATAGACTATGATTGGTGTGAAGATGATTATCAACAG 1320

Qy 1321 GTGCTGCTACTCCCGGAAACATCACAGTTTACCAAGTACTGGACAGAAATCAAAATGGGGTG 1380

Db 1321 GTGCTGCTACTCCCGGAAACATCACAGTTTACCAAGTACTGGACAGAAATCAAAATGGGGTG 1380

Qy 1381 AATCTAAATCATTAATCTCAGCCTTATGAGAACACCTGCAAAATAAATTAAGAGAACAAAG 1440

Db 1381 AATCTAAATCATTAATCTCAGCCTTATGAGAACACCTGCAAAATAAATTAAGAGAACAAAG 1440

Qy 1441 AAAATGTATATCTCTTAAGTCTGCTGTAAGAAATCAAGAGTACTTATGTTTCTCGAGC 1500

Db 1441 AAAATGTATATCTCTTAAGTCTGCTGTAAGAAATCAAGAGTACTTATGTTTCTCGAGC 1500

Qy 1501 CAAAGACTCCAGTTAATAAGAACCCAGCATAAGAGAGAAATACTCTACAGCCCAAAATCGTT 1560

Db 1501 CAAAGACTCCAGTTAATAAGAACCCAGCATAAGAGAGAAATACTCTACAGCCCAAAATCGTT 1560

Qy 1561 ACACATACCCCTCAAAAGCTAGAAACGAGTGCCTGAAAGAACTCCAATTAATAATACAG 1620

Db 1561 ACACATACCCCTCAAAAGCTAGAAACGAGTGCCTGAAAGAACTCCAATTAATAATACAG 1620

Qy 1621 TAAATTTCAACAGGAAACAGACAGTTAATGACAGGTCTCATTAGCCCTGAGAGCGGTGCC 1680

Db 1621 TAAATTTCAACAGGAAACAGACAGTTAATGACAGGTCTCATTAGCCCTGAGAGCGGTGCC 1680

Qy 1681 GCTCAGTGGAAATTTGGATCTCAACCAAGCACATATGAGGAGACTCCAAAGAAAGAGGAG 1740

Db 1681 GCTCAGTGGAAATTTGGATCTCAACCAAGCACATATGAGGAGACTCCAAAGAAAGAGGAG 1740

Qy 1741 CCAAGTGTTTGGAGCCTTGAAGGGGTTGGATAAGGTTATCACTGTGCTCACAGGA 1800

Db 1741 CCAAGTGTTTGGAGCCTTGAAGGGGTTGGATAAGGTTATCACTGTGCTCACAGGA 1800

Qy 1801 GCAAAAGGAGGGTCTGCCAGAGACGGGCCAGAGACTAAAGCTTCACTATAATGTGA 1860

Db 1801 GCAAAAGGAGGGTCTGCCAGAGACGGGCCAGAGACTAAAGCTTCACTATAATGTGA 1860

Qy 1861 CTACAACTAGATPAGTGAATCCAGATCAACTGTTTGAATGAATAATATGTCTATTCTTCCAA 1920

Db 1861 CTACAACTAGATPAGTGAATCCAGATCAACTGTTTGAATGAATAATATGTCTATTCTTCCAA 1920

Qy 1921 AGAAGCATGTGTGACTTTGTATCAAAAAGGGTTATACACTGAAGTGTCAAAACCCGATGTG 1980

Db 1921 AGAAGCATGTGTGACTTTGTATCAAAAAGGGTTATACACTGAAGTGTCAAAACCCGATGTG 1980

Qy 1981 TTGGGAAAGTGCAATGCAATTTGAATTTAGAAGTGTGCCAGCTTCAAAAACCCGATGTG 2040

Db 1981 TTGGGAAAGTGCAATGCAATTTGAATTTAGAAGTGTGCCAGCTTCAAAAACCCGATGTG 2040

Qy 2041 TGGGTATCAGGAGCGAGCGCTTAAAGCGATGCTTGGGTTTACAAAAGATTAGTGGAA 2100

Db 2041 TGGGTATCAGGAGCGAGCGCTTAAAGCGATGCTTGGGTTTACAAAAGATTAGTGGAA 2100

Qy 2101 ACATCCTATCTAGCTGCAAGGTATAATTTGATGATTCCTTCATCTCCGCGATGAGTGTG 2160

Db 2101 ACATCCTATCTAGCTGCAAGGTATAATTTGATGATTCCTTCATCTCCGCGATGAGTGTG 2160

Qy 2161 GGTGTGATACAGCCTACATAAAGCTGTTATGATCGCTTTGATTTTAAAGTTTCATTTGGAA 2220

Db 2161 GGTGTGATACAGCCTACATAAAGCTGTTATGATCGCTTTGATTTTAAAGTTTCATTTGGAA 2220

Qy 2221 CTACCAACTTGTGTTCTTAAAGAGCTATCTTAAAGACCAATATCTCTTTGTTTTTAAACAAA 2280

Db 2221 CTACCAACTTGTGTTCTTAAAGAGCTATCTTAAAGACCAATATCTCTTTGTTTTTAAACAAA 2280

Qy 2281 GATATATTTTGTGTATGAATCTAAATCAAGCCCATCTGCTATATGTTACTGCTCTTTT 2340

Db 2281 GATATATTTTGTGTATGAATCTAAATCAAGCCCATCTGCTATATGTTACTGCTCTTTT 2340

Qy 2341 TAATCATGTGTTTTGTATATTAATAATTTGTGACTTTCTTAGATTCACCTTCCATATGTG 2400

Db 2341 TAATCATGTGTTTTGTATATTAATAATTTGTGACTTTCTTAGATTCACCTTCCATATGTG 2400

Qy 2401 AATGTAAGCTCTTAACATATGCTCTTTGTAAATGTGTAATTTCTTGAATAAAACCAT 2460

Db 2401 AATGTAAGCTCTTAACATATGCTCTTTGTAAATGTGTAATTTCTTCTGAAATAAAACCAT 2460

Qy 2461 TTGTGAATAT 2470

Db 2461 TTGTGAATAT 2470

RESULT 13

US-10-756-149-2259  
; Sequence 2259, Application US/10756149  
; Publication NO. US20050181375A1  
; GENERAL INFORMATION:  
; APPLICANT: Aziz, Natasha  
; APPLICANT: Zlotnik, Albert  
; TITLE OF INVENTION: NOVEL METHODS OF DIAGNOSIS OF METASTATIC CANCER, COMPOSITIONS AND  
; TITLE OF INVENTION: METHODS OF SCREENING FOR MODULATORS OF METASTATIC CANCER  
; FILE REFERENCE: file  
; CURRENT APPLICATION NUMBER: US/10/756,149  
; CURRENT FILING DATE: 2004-01-12  
; NUMBER OF SEQ ID NOS: 5818  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 2259  
; LENGTH: 2470  
; TYPE: DNA  
; ORGANISM: Homo Sapiens  
US-10-756-149-2259

Query Match 100.0%; Score 2470; DB 10; Length 2470;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 2470; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TTGGCGGGCGGAAAGCGGCCACAAACCGCGCATCGAAAAGATTCTTAGGAACGCCGTACCA 60

Db 1 TTGGCGGGCGGAAAGCGGCCACAAACCGCGCATCGAAAAGATTCTTAGGAACGCCGTACCA 60

Qy 61 GCCCGCTCTCTCAGAACAGCAGCGCCCTGTCTTCTGTGGCGCGCGCTCAGCGGTGCC 120

Db 61 GCCCGCTCTCTCAGAACAGCAGCGCCCTGTCTTCTGTGGCGCGCGCTCAGCGGTGCC 120

Qy 121 TCCGCCCTCAGTCTTTTCTTAATTCAAATAAATTCGAAGAGACTATGAAGATT 180

Db 121 TCCGCCCTCAGTCTTTTCTTAATTCAAATAAATTCGAAGAGACTATGAAGATT 180

Qy 181 ATGATGAACCTTCTCAAAATATTAATGAATTAACATGAATACTATGGGACAGGTGGCTTGCAA 240

Db 181 ATGATGAACCTTCTCAAAATATTAATGAATTAACATGAATACTATGGGACAGGTGGCTTGCAA 240

Qy 241 AGGTCAAACTTGCCTGCCATATCTTACTTGGAGAGATGGTAGCTATAAAAAATCATGGATA 300

Db 241 AGGTCAAACTTGCCTGCCATATCTTACTTGGAGAGATGGTAGCTATAAAAAATCATGGATA 300

Qy 301 AAAAAACACTAGGAGTGTATTTGCCCGGATCAAAACCGAGATTGAGGCTTGAAGACC 360

Db 301 AAAAAACACTAGGAGTGTATTTGCCCGGATCAAAACCGAGATTGAGGCTTGAAGACC 360

Qy 361 TGAGACATCAGCATATATGTCAACTCTACATGTGTAGAGACAGCAACAAAAATTCA 420

Db 361 TGAGACATCAGCATATATGTCAACTCTACATGTGTAGAGACAGCAACAAAAATTCA 420

Qy 421 TGGTCTTGAAGTACTGCCCTCGAGAGAGCTGTTTGAATATATATTTCCAGGATCGCC 480

Db 421 TGGTCTTGAAGTACTGCCCTCGAGAGAGCTGTTTGAATATATATTTCCAGGATCGCC 480

Qy 481 TGTCAAGAGAGAGACCCGGGTTGCTTCCGTCAGATAGTATCTGCTGTGTTGCTTATGTGC 540

Db 481 TGTCAAGAGAGAGACCCGGGTTGCTTCCGTCAGATAGTATCTGCTGTGTTGCTTATGTGC 540



; CURRENT FILING DATE: 2004-02-04									
; NUMBER OF SEQ ID NOS: 48640									
; SOFTWARE: PatentIn version 3.2									
; SEQ ID NO 23									
; LENGTH: 2470									
; TYPE: DNA									
; ORGANISM: Homo sapiens									
US-10-770-726-23									
Query Match 100.0%; Score 2470; DB 10; Length 2470;									
Best Local Similarity 100.0%; Pred. No. 0;									
Matches 2470; Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
QY	1	TTGGCGGGCGGAAGCGGCCACACCCCGGATCGAAAGATTCTTAGGAACGCGTACCA	60						
DB	1	TTGGCGGGCGGAAGCGGCCACACCCCGGATCGAAAGATTCTTAGGAACGCGTACCA	60						
QY	61	GC CGCGTCTCTCAGGACAGCAGGCCCTCTCTTCTGTGCGGCGCGCTCAGCCGTGCC	120						
DB	61	GC CGCGTCTCTCAGGACAGCAGGCCCTCTCTTCTGTGCGGCGCGCTCAGCCGTGCC	120						
QY	121	TCGCGCCCTCAGGTTCTTTTCTTAATTCBAATAAACTTGCAGAGGACTATGAAGATT	180						
DB	121	TCGCGCCCTCAGGTTCTTTTCTTAATTCBAATAAACTTGCAGAGGACTATGAAGATT	180						
QY	181	ATGATGAACCTTCTCAATATATGAATTACATGAATTAATTCGAAGAGTTTGGCAA	240						
DB	181	ATGATGAACCTTCTCAATATATGAATTACATGAATTAATTCGAAGAGTTTGGCAA	240						
QY	241	AGGTCAAACTTGGCTGCGCATATCTTTACTGAGAGAGTGTAGTATAAAATCATGGATA	300						
DB	241	AGGTCAAACTTGGCTGCCATATCTTTACTGAGAGAGTGTAGTATAAAATCATGGATA	300						
QY	301	AAAACACACTAGGAGTGATTTGCCCGGATCAAAACGGAGATTGAGCCCTTGAAGACC	360						
DB	301	AAAACACACTAGGAGTGATTTGCCCGGATCAAAACGGAGATTGAGCCCTTGAAGACC	360						
QY	361	TGAGACATCAGCATATATGTCAACTCTACCATGTCTAGAGACAGCCCAAAATATTCA	420						
DB	361	TGAGACATCAGCATATATGTCAACTCTACCATGTCTAGAGACAGCCCAAAATATTCA	420						
QY	421	TGGTCTTGTAGTACTGCCCTGGAGGAGCTGTTTGACTATATAATTTCCAGGATCGCC	480						
DB	421	TGGTCTTGTAGTACTGCCCTGGAGGAGCTGTTTGACTATATAATTTCCAGGATCGCC	480						
QY	481	TGTCAGAAGAGGAGACCCCGGTTGTCTCCGTGAGATAGTATCTGTGCTTATGTGC	540						
DB	481	TGTCAGAAGAGGAGACCCCGGTTGTCTCCGTGAGATAGTATCTGTGCTTATGTGC	540						
QY	541	ACAGCCAGGGCTATGTCTCAGGGACCTCAAGCCAGAAAATTTGCTGTTGATGAATATC	600						
DB	541	ACAGCCAGGGCTATGTCTCAGGGACCTCAAGCCAGAAAATTTGCTGTTGATGAATATC	600						
QY	601	ATAAATTAAGCTGATTGACTTTTGGTCTGTGTCAAACCCAAAGGGTAAACAAGATTACC	660						
DB	601	ATAAATTAAGCTGATTGACTTTTGGTCTGTGTCAAACCCAAAGGGTAAACAAGATTACC	660						
QY	661	ATCTACAGACATGCTGTGGAGTCTGCTTATGACGACCTGATGTTAATCAAGGCAAT	720						
DB	661	ATCTACAGACATGCTGTGGAGTCTGCTTATGACGACCTGATGTTAATCAAGGCAAT	720						
QY	721	CATATCTTGGATCAGAGGACAGATGTTTGGAGCATGGGCATCTGTTATATGTTCTTATGT	780						
DB	721	CATATCTTGGATCAGAGGACAGATGTTTGGAGCATGGGCATCTGTTATATGTTCTTATGT	780						
QY	781	GTGGATTCTCACCATTGTGATGATAATGTAATGGCTTTATACAAGAGATTATGAGAG	840						
DB	781	GTGGATTCTCACCATTGTGATGATAATGTAATGGCTTTATACAAGAGATTATGAGAG	840						
QY	841	GAATAATATGATGTTCCCAAGTGGCTCTCTCCAGTAGCATTTCTTCTTCAACAATGC	900						
DB	841	GAATAATATGATGTTCCCAAGTGGCTCTCTCCAGTAGCATTTCTTCTTCAACAATGC	900						
QY	901	TGCAGGTGACCCCAAGAAACGGATTTCTATGAAAAATCTATTGAACCATCCTCTGGATCA	960						
DB	901	TGCAGGTGACCCCAAGAAACGGATTTCTATGAAAAATCTATTGAACCATCCTCTGGATCA	960						
QY	961	TGCAAGATTACAACATATCCTGTTGAGTGCCAAAGCAAGAAATCCTTTATTACCTCGATG	1020						
DB	961	TGCAAGATTACAACATATCCTGTTGAGTGCCAAAGCAAGAAATCCTTTATTACCTCGATG	1020						
QY	1021	ATGATTGGGTAAACAGAACTTTCTGTACATCAAGAAAAACAGGAAAAACAAATGGAGATT	1080						
DB	1021	ATGATTGGGTAAACAGAACTTTCTGTACATCAAGAAAAACAGGAAAAACAAATGGAGATT	1080						
QY	1081	TAAATTTCACTGTGGCAGTATGATCACCTCACGGCTACCTATCTTCTGTCTTAGCCAAAG	1140						
DB	1081	TAAATTTCACTGTGGCAGTATGATCACCTCACGGCTACCTATCTTCTGTCTTAGCCAAAG	1140						
QY	1141	AGGCTCGGGGAAAAACGAGTTCTTTAAGGCTTTCTTTCTCTCTGTGCAAGCCAGTG	1200						
DB	1141	AGGCTCGGGGAAAAACGAGTTCTTTAAGGCTTTCTTTCTCTCTGTGCAAGCCAGTG	1200						
QY	1201	CTACCCCTTTCACAGACATCAAGTCAAAATAATTTGGAGTCTGGAAGATGTGACCGAAGTG	1260						
DB	1201	CTACCCCTTTCACAGACATCAAGTCAAAATAATTTGGAGTCTGGAAGATGTGACCGAAGTG	1260						
QY	1261	ATAAAAAATTATGTGGCGGATTAATAGACTATGATTTGGTGTGAAGATGATTTATCAACAG	1320						
DB	1261	ATAAAAAATTATGTGGCGGATTAATAGACTATGATTTGGTGTGAAGATGATTTATCAACAG	1320						
QY	1321	GTGCTGCTACTCCCGAATCATCAGTTTACAGTTTACAGTACTGGAACAGATCAAAATGGGGTG	1380						
DB	1321	GTGCTGCTACTCCCGAATCATCAGTTTACAGTTTACAGTACTGGAACAGATCAAAATGGGGTG	1380						
QY	1381	AATCTAAATCATTAATCTCCAGCTTATGAGAACACCTGCAAAATAAATTAAGAACAAAG	1440						
DB	1381	AATCTAAATCATTAATCTCCAGCTTATGAGAACACCTGCAAAATAAATTAAGAACAAAG	1440						
QY	1441	AAAAATGATATATCTCCTAAGTCTGTGTAAGAAATGAAGAGTACTTTATGTTCTCTGAGC	1500						
DB	1441	AAAAATGATATATCTCCTAAGTCTGTGTAAGAAATGAAGAGTACTTTATGTTCTCTGAGC	1500						
QY	1501	CAAAGACTCCAGTTTAATAAGAAACAGATTAAGAGAGAAATCTACTACGCCAAATCGTT	1560						
DB	1501	CAAAGACTCCAGTTTAATAAGAAACAGATTAAGAGAGAAATCTACTACGCCAAATCGTT	1560						
QY	1561	ACACTACACCTCAAAGCTAGAAAACAGTCTGCTGAAAGAAATCTCAAATTAATAACAG	1620						
DB	1561	ACACTACACCTCAAAGCTAGAAAACAGTCTGCTGAAAGAAATCTCAAATTAATAACAG	1620						
QY	1621	TAAATTCACAGGAAACAGATTTAATGACAGGTGTCAATTAGCCCTGAGAGCGGTGCC	1680						
DB	1621	TAAATTCACAGGAAACAGATTTAATGACAGGTGTCAATTAGCCCTGAGAGCGGTGCC	1680						
QY	1681	GCTCAGTGAATTTGGATCTCAACCAAGCAGATATGAGAGAGATCTCCAAAAGAAAGGAG	1740						
DB	1681	GCTCAGTGAATTTGGATCTCAACCAAGCAGATATGAGAGAGATCTCCAAAAGAAAGGAG	1740						
QY	1741	CAAAAGTGTGAGGACCTTTGAAAGGGGTGTGATTAAGGTATCACTGTGCTCACCAGGA	1800						
DB	1741	CAAAAGTGTGAGGACCTTTGAAAGGGGTGTGATTAAGGTATCACTGTGCTCACCAGGA	1800						
QY	1801	GCAAAAGGAGGGTCTCGCCAGAGACGGGCCCAAGAGACTTAAAGCTTCACTATAATGTGA	1860						
DB	1801	GCAAAAGGAGGGTCTCGCCAGAGACGGGCCCAAGAGACTTAAAGCTTCACTATAATGTGA	1860						
QY	1861	CTACAACTAGATTAGTGAATCCAGATCAACTGTTGAAATGAATTAATGTTCTTTCCAA	1920						
DB	1861	CTACAACTAGATTAGTGAATCCAGATCAACTGTTGAAATGAATTAATGTTCTTTCCAA	1920						
QY	1921	AGAAGCATGTTGACTTTTGTACAAAAGGGTTATACACTGAAAGTGTCAAAACACAGTCAAGAT	1980						
DB	1921	AGAAGCATGTTGACTTTTGTACAAAAGGGTTATACACTGAAAGTGTCAAAACACAGTCAAGAT	1980						
QY	1981	TTGGGAAAGTGACAAATTTGCAATTTGAAGTGTGCCAGCTTCAAAAACCCCGATGTGG	2040						

1981	Db		TTGGGAAGTGCACAAATGCAATTTTGAATTTAGAAAGTGTCCAGCTTCAAAAACCCCGATGTGG	2040
2041	Qy		TGGGTATCAGGAGGACAGCGGCTTAAGGGCGATGCTGGGTTTACAAAAGATTAGTGAAG	2100
2041	Db		TGGGTATCAGGAGGACAGCGGCTTAAGGGCGATGCTGGGTTTACAAAAGATTAGTGAAG	2100
2101	Qy		ACATCCTATCTAGCTGCAAGGTAAATPGATGGAACTTTCCATCCTCCCGATGAGTGTG	2160
2101	Db		ACATCCTATCTAGCTGCAAGGTAAATPGATGGAACTTTCCATCCTCCCGATGAGTGTG	2160
2161	Qy		GGTGTGATACAGCCTACATAAAGACTGTTATGATCGCTTTTGATTTTAAAGTTCATTGGAA	2220
2161	Db		GGTGTGATACAGCCTACATAAAGACTGTTATGATCGCTTTTGATTTTAAAGTTCATTGGAA	2220
2221	Qy		CTACCAACTTGTGTTTCTAAAGAGCTATCTTAAGACCAATATCTCTTTCGTTTTTAAACAAAA	2280
2221	Db		CTACCAACTTGTGTTTCTAAAGAGCTATCTTAAGAGCTATCTCTTTCGTTTTTAAACAAAA	2280
2281	Qy		GATATTAATTTTGTGTATGAATCTAAATCAAGCCCATCTGTCAATPATGTTACTGTCTTTTT	2340
2281	Db		GATATTAATTTTGTGTATGAATCTAAATCAAGCCCATCTGTCAATPATGTTACTGTCTTTTT	2340
2341	Qy		TAATCATGTGGTTTTGTGTATTAATTAATGTTGACTTTCTTAGATTCACATCCATATGTG	2400
2341	Db		TAATCATGTGGTTTTGTGTATTAATTAATGTTGACTTTCTTAGATTCACATCCATATGTG	2400
2401	Qy		AAATGAAGCTCTTAAACTATGCTCTTTTGTAAATGTGTAAATTTCTTTCTGAAATAAAACCAT	2460
2401	Db		AAATGAAGCTCTTAAACTATGCTCTTTTGTAAATGTGTAAATTTCTTTCTGAAATAAAACCAT	2460
2461	Qy		TTGTGGAATAT	2470
2461	Db		TTGTGGAATAT	2470

RESULT 15

```

US-10-960-414-132
; Sequence 132, Application US/10960414
; Publication No. US20060074565A1
; GENERAL INFORMATION:
; APPLICANT: MILLER, LANCE D.
; APPLICANT: GEORGE, JOSH V.
; APPLICANT: VEGA, VINSENSIUS B.
; TITLE OF INVENTION: METHODS, SYSTEMS, AND COMPOSITIONS FOR CLASSIFICATION,
; TITLE OF INVENTION: PROGNOSIS, AND DIAGNOSIS OF CANCERS
; FILE REFERENCE: 38271-76067
; CURRENT APPLICATION NUMBER: US/10/960,414
; CURRENT FILING DATE: 2004-10-06
; NUMBER OF SEQ ID NOS: 500
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 132
; LENGTH: 2470
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-960-414-132

```

QY	181	ATGATGAAC	TTCTCAAATATATATGAATTA	TACATGA	AACTATTTGGGACAGGTGGCTTTTGC	CAA	240	
DB	181	ATGATGAAC	TTCTCAAATATATATGAATTA	TACATGA	AACTATTTGGGACAGGTGGCTTTTGC	CAA	240	
QY	241	AGGTCAAAC	TTGCGTCGCATATCCTTACT	TGGAGAGATGGT	AGCTATATAAAATCAT	TGATCA	300	
DB	241	AGGTCAAAC	TTGCGTCGCATATCCTTACT	TGGAGAGATGGT	AGCTATATAAAATCAT	TGATCA	300	
QY	301	AAAAACA	CTAGGAGATGATTTGGCCCGGAT	CAAAACGGAGAT	TGAGGCGCTTGAAGA	CAACC	360	
DB	301	AAAAACA	CTAGGAGATGATTTGGCCCGGAT	CAAAACGGAGAT	TGAGGCGCTTGNAGNA	CAACC	360	
QY	361	TGAGACAT	CAGCATATATATGTCAA	CTCTCAATGTGCT	TAGAGACAGCCAA	CAAAATATTC	420	
DB	361	TGAGACAT	CAGCATATATATGTCAA	CTCTCAATGTGCT	TAGAGACAGCCAA	CAAAATATTC	420	
QY	421	TGGTTCTTG	AGTACTGCGCCCTGGAGGAG	CTGTTTGACTATATA	TTTCCCGAGGATCGCC		480	
DB	421	TGGTTCTTG	AGTACTGCGCCCTGGAGGAG	CTGTTTGACTATATA	TTTCCCGAGGATCGCC		480	
QY	481	TGTCAGA	AGAGGAGACCCGGGTGTCT	CTCCGT	CAGATAGTATCTGCTGTTGCTT	TATGTGC	540	
DB	481	TGTCAGA	AGAGGAGACCCGGGTGTCT	CTCCGT	CAGATAGTATCTGCTGTTGCTT	TATGTGC	540	
QY	541	ACGCCAG	GGGCTATGTCTCACAGG	CACTCAAGCCAGAAA	ATTTGCTGTTTGATCA	ATATC	600	
DB	541	ACGCCAG	GGGCTATGTCTCACAGG	CACTCAAGCCAGAAA	ATTTGCTGTTTGATCA	ATATC	600	
QY	601	ATAAATTA	AGCTGATTTGACTTTGGTCT	CTGTGCA	AAACCCAGGGGTAA	CAAGGATTA	660	
DB	601	ATAAATTA	AGCTGATTTGACTTTGGTCT	CTGTGCA	AAACCCAGGGGTAA	CAAGGATTA	660	
QY	661	ATCTAC	CAGACATGCTGTGGAG	CTGCGTTATGC	CAGCACTGAGTTTAAT	ACAAGGCCAAAT	720	
DB	661	ATCTAC	CAGACATGCTGTGGAG	CTGCGTTATGC	CAGCACTGAGTTTAAT	ACAAGGCCAAAT	720	
QY	721	CATATCT	TGGATCAGAGG	CAGATGTTTGGAG	CAATGGGCATATCTGTTATAT	TGTTCTTATGT	780	
DB	721	CATATCT	TGGATCAGAGG	CAGATGTTTGGAG	CAATGGGCATATCTGTTATAT	TGTTCTTATGT	780	
QY	781	GTGGATTT	CTACCAATTTGATGATNA	TGTATGGCTTTTAT	CAAGAGAGATTTAT	TGAGAG	840	
DB	781	GTGGATTT	CTACCAATTTGATGATNA	TGTATGGCTTTTAT	CAAGAGAGATTTAT	TGAGAG	840	
QY	841	GAAATAT	GATGTTTCCCAAGTGG	CTCTCC	CAGTAGCATTTCTGCTTCTT	CTTCAACAATGC	900	
DB	841	GAAATAT	GATGTTTCCCAAGTGG	CTCTCC	CAGTAGCATTTCTGCTTCTT	CTTCAACAATGC	900	
QY	901	TGCAAGT	TGGACCCAAAGAAA	CGGATTTCTATG	AAAAATCTATTGAA	CCATCCCTGGATCA	960	
DB	901	TGCAAGT	TGGACCCAAAGAAA	CGGATTTCTATG	AAAAATCTATTGAA	CCATCCCTGGATCA	960	
QY	961	TGCNAG	ATTAACA	CTATCCTGTTGAGT	TGGCAAGCAAGATCC	CTTTTATTCACCTCGATG	1020	
DB	961	TGCNAG	ATTAACA	CTATCCTGTTGAGT	TGGCAAGCAAGATCC	CTTTTATTCACCTCGATG	1020	
QY	1021	ATGATT	TCGGTAA	CAGAAC	TTTCTGTGAT	CATCACAGAA	CAACAGGCAAA	1080
DB	1021	ATGATT	TCGGTAA	CAGAAC	TTTCTGTGAT	CATCACAGAA	CAACAGGCAAA	1080
QY	1081	TAAATTT	CACTGTGG	CAGTATGAT	CACCTC	CAGGCTACCTATCTTCT	TCTAGCCAA	1140
DB	1081	TAAATTT	CACTGTGG	CAGTATGAT	CACCTC	CAGGCTACCTATCTTCT	TCTAGCCAA	1140
QY	1141	AGGCT	CGGGAAAA	CCAGTTTCGTTT	TAAAGGCTTTTCTT	CTTCTCTCTGTG	CNACAGCCAGTG	1200
DB	1141	AGGCT	CGGGAAAA	CCAGTTTCGTTT	TAAAGGCTTTTCTT	CTTCTCTCTGTG	CNACAGCCAGTG	1200
QY	1201	CTACCC	CAATTC	CAGACATCA	AGTCAAA	TAAATTTGGAGTCT	TGGAAGATGTGACCGCAAGTG	1260
DB	1201	CTACCC	CAATTC	CAGACATCA	AGTCAAA	TAAATTTGGAGTCT	TGGAAGATGTGACCGCAAGTG	1260
QY	1261	ATAAA	ATTA	TGTGGCGG	ATTAATAG	ATGATTTGGTGT	TGAAAGATGATTTATCAACAG	1320





GenCore version 5.1.9  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 21, 2006, 19:43:24 ; Search time 560 Seconds  
(without alignments)  
8770.613 Million cell updates/sec

Title: US-10-656-598-1  
Perfect score: 2470  
Sequence: 1 ttggcggggcgaagcgcca.....ataaacatttgtaatat 2470

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 2477703 seqs, 994239549 residues

Total number of hits satisfying chosen parameters: 4955406

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : 1: /EMC\_Celerra\_SID33/ptodata/2/pubnpna/US09\_NEW\_PUB.seq.\*

- 2: /EMC\_Celerra\_SID33/ptodata/2/pubnpna/US06\_NEW\_PUB.seq.\*
- 3: /EMC\_Celerra\_SID33/ptodata/2/pubnpna/US07\_NEW\_PUB.seq.\*
- 4: /EMC\_Celerra\_SID33/ptodata/2/pubnpna/US08\_NEW\_PUB.seq.\*
- 5: /EMC\_Celerra\_SID33/ptodata/2/pubnpna/PCT\_NEW\_PUB.seq.\*
- 6: /EMC\_Celerra\_SID33/ptodata/2/pubnpna/US10\_NEW\_PUB.seq.\*
- 7: /EMC\_Celerra\_SID33/ptodata/2/pubnpna/US11\_NEW\_PUB.seq.\*
- 8: /EMC\_Celerra\_SID33/ptodata/2/pubnpna/US11\_NEW\_PUB.seq.\*
- 9: /EMC\_Celerra\_SID33/ptodata/2/pubnpna/US11\_NEW\_PUB.seq.\*
- 10: /EMC\_Celerra\_SID33/ptodata/2/pubnpna/US60\_NEW\_PUB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	937	37.9	1000	8	US-11-266-748A-222300
2	937	37.9	1000	8	US-11-266-748A-228390
3	937	37.9	1000	8	US-11-266-748A-339819
4	937	37.9	1000	8	US-11-266-748A-339255
5	937	37.9	1000	8	US-11-266-748A-470301
6	901.6	36.5	926	8	US-11-266-748A-189434
7	901.6	36.5	926	8	US-11-266-748A-194943
8	791.2	32.0	1050	8	US-11-266-748A-262797
9	791.2	32.0	1050	8	US-11-266-748A-323314
10	717.6	29.1	758	8	US-11-266-748A-9105
11	685.4	27.7	948	8	US-11-266-748A-189435
12	685.4	27.7	948	8	US-11-266-748A-194944
13	590.6	23.9	728	8	US-11-266-748A-189433
14	514	20.8	520	8	US-11-266-748A-44240
15	208.4	8.4	1863	9	US-11-145-471-52
16	206.8	8.4	5093	8	US-11-266-748A-23698
17	202.4	8.2	2213	6	US-10-449-902-25149
18	202.4	8.2	2246	6	US-10-449-902-22550
19	197.8	8.0	2089	9	US-11-056-355B-48678
20	197.8	8.0	2089	9	US-11-056-355B-49794
21	196	7.9	2927	6	US-10-449-902-18825
22	192.8	7.8	2476	9	US-11-218-305-16612
23	192.6	7.8	1326	7	US-11-299-286-72

24	192.6	7.8	1569	9	US-11-056-355B-89480	Sequence 89480, A
25	192.6	7.8	1569	9	US-11-056-355B-93236	Sequence 93236, A
26	192.6	7.8	1585	9	US-11-056-355B-47225	Sequence 47225, A
27	192.6	7.8	1724	9	US-11-056-355B-89484	Sequence 89484, A
28	192.6	7.8	1724	9	US-11-056-355B-93240	Sequence 93240, A
29	190	7.7	809	8	US-11-266-748A-14709	Sequence 14709, A
30	187.6	7.6	1229	9	US-11-056-355B-18719	Sequence 18719, A
31	187.6	7.6	2126	9	US-11-218-305-3767	Sequence 3767, Ap
32	186	7.5	1990	6	US-10-953-349-10926	Sequence 10926, A
33	184.8	7.5	1539	7	US-11-299-286-2657	Sequence 2657, Ap
34	182.8	7.4	2043	9	US-11-218-305-3768	Sequence 3768, Ap
35	182	7.4	1909	6	US-10-953-349-25259	Sequence 25259, A
36	182	7.4	1909	9	US-11-056-355B-67583	Sequence 67583, A
c	182	7.4	2073	9	US-11-218-305-24784	Sequence 24784, A
37	180.8	7.3	2108	6	US-10-449-902-22705	Sequence 22705, A
38	179	7.2	2017	6	US-10-612-783-997	Sequence 997, App
40	178.8	7.2	1642	8	US-11-266-748A-252242	Sequence 252242, A
41	178.8	7.2	1642	8	US-11-266-748A-275764	Sequence 275764, A
c	178.8	7.2	1642	8	US-11-266-748A-312759	Sequence 312759, A
42	178.4	7.2	1143	8	US-11-216-545-7035	Sequence 7035, Ap
44	178.4	7.2	1341	9	US-11-040-005-1	Sequence 1, Appli
45	178.4	7.2	1754	9	US-11-056-355B-47475	Sequence 47475, A

#### ALIGNMENTS

RESULT 1  
US-11-266-748A-222300  
; Sequence 222300, Application US/11266748A  
; Publication No. US20060134663A1  
; GENERAL INFORMATION:  
; APPLICANT: Harkin, Paul  
; APPLICANT: Johnston, Patrick  
; APPLICANT: Mulligan, Karl  
; TITLE OF INVENTION: Transcriptome Microarray Technology and  
; TITLE OF INVENTION: Methods of Using the Same  
; FILE REFERENCE: 55815-0102 (319189)  
; CURRENT APPLICATION NUMBER: US/11/266,748A  
; CURRENT FILING DATE: 2005-11-03  
; PRIOR APPLICATION NUMBER: EP 04105479.2  
; PRIOR FILING DATE: 2004-11-03  
; PRIOR APPLICATION NUMBER: EP 04105482.6  
; PRIOR FILING DATE: 2004-11-03  
; PRIOR APPLICATION NUMBER: EP 04105483.4  
; PRIOR FILING DATE: 2004-11-03  
; PRIOR APPLICATION NUMBER: EP 04105507.0  
; PRIOR FILING DATE: 2004-11-03  
; PRIOR APPLICATION NUMBER: EP 04105485.9  
; PRIOR FILING DATE: 2004-11-03  
; PRIOR APPLICATION NUMBER: EP 04105484.2  
; PRIOR FILING DATE: 2004-11-03  
; PRIOR APPLICATION NUMBER: US 60/662,276  
; PRIOR FILING DATE: 2005-03-14  
; PRIOR APPLICATION NUMBER: US 60/700,293  
; PRIOR FILING DATE: 2005-07-18  
; NUMBER OF SEQ ID NOS: 483996  
; SOFTWARE: PatentIn version 3.3  
; SEQ ID NO 222300  
; LENGTH: 1000  
; TYPE: DNA  
; ORGANISM: Homo Sapiens  
US-11-266-748A-222300

Query Match 37.9%; Score 937; DB 8; Length 1000;  
Best Local Similarity 100.0%; Pred. No. 2.6e-236;  
Matches 937; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1534 GAGAAATCTACTACGCGCAATCGTTTACCTACACCTCTCAAAAGCTAGAACCACTGCGC 1593  
1 GAGAAATCTACTACGCGCAATCGTTTACCTACACCTCTCAAAAGCTAGAACCACTGCGC 60  
QY 1594 TGAAGAAACTCCAATTAAATACCAGTAAATTCACAGGAAACAGCAAGTTAATGACAG 1653

Db 61 TGAAGAAATCCCAATTAATAATACCAAGTAATTAATGACAG 120  
Qy 1654 GTGTCAATAGCCCTGAGAGCGGTGCGCTCAGTGGAAATGGATCTCAACAGACATA 1713  
Db 121 GTGTCAATAGCCCTGAGAGCGGTGCGCTCAGTGGAAATGGATCTCAACAGACATA 180  
Qy 1714 TGAAGAGACTCCAAAGAAAGAGGAGCCAAAGTGTGTTGGAGCCCTTGAAGGGGGTTGG 1773  
Db 181 TGAAGAGACTCCAAAGAAAGAGGAGCCAAAGTGTGTTGGAGCCCTTGAAGGGGGTTGG 240  
Qy 1774 ATAAGGTATCACTGTGCTCACCAGAGCAAAAGAGGGTCTGCCAGAGAGCGGCCCA 1833  
Db 241 ATAAGGTATCACTGTGCTCACCAGAGCAAAAGAGGGTCTGCCAGAGAGCGGCCCA 300  
Qy 1834 GAGACTTAAGCTTCACTATAATGTGACTACACTAGATAGTGAATCCAGATCAACTGT 1893  
Db 301 GAAGACTTAAGCTTCACTATAATGTGACTACACTAGATAGTGAATCCAGATCAACTGT 360  
Qy 1894 TGAATGAATAATGTCTATTCTTCCAAAGAGCATGTTGACTTTGTACAAAAGGGTTATA 1953  
Db 361 TGAATGAATAATGTCTATTCTTCCAAAGAGCATGTTGACTTTGTACAAAAGGGTTATA 420  
Qy 1954 CACTGAAGTGTCAAAACACAGTCAAGATTTTGGGAAAGTGACAATGCAATTTGAATTAAG 2013  
Db 421 CACTGAAGTGTCAAAACACAGTCAAGATTTTGGGAAAGTGACAATGCAATTTGAATTAAG 480  
Qy 2014 TGTGCCAGCTTCAAAAACCCGATGTGTGGGTATCAGGAGGCGCGCTTAAGGGCGATG 2073  
Db 481 TGTGCCAGCTTCAAAAACCCGATGTGTGGGTATCAGGAGGCGCGCTTAAGGGCGATG 540  
Qy 2074 CTGGGGTTTCAAAAGATTAAGTGAAGACATCTCTAGCTGCAAGGTATAATTCATGG 2133  
Db 541 CTGGGGTTTCAAAAGATTAAGTGAAGACATCTCTAGCTGCAAGGTATAATTCATGG 600  
Qy 2134 ATTCTTCCATCTGCGCGATGAGTGTGGGTGTGATACAGCCTACATAAAGACTTTATGA 2193  
Db 601 ATTCTTCCATCTGCGCGATGAGTGTGGGTGTGATACAGCCTACATAAAGACTTTATGA 660  
Qy 2194 TCGCTTTGATTTTAAAGTTCATTTGGAACTACCAACTTGTGTTCTTAAAGAGCTATCTTAAGA 2253  
Db 661 TCGCTTTGATTTTAAAGTTCATTTGGAACTACCAACTTGTGTTCTTAAAGAGCTATCTTAAGA 720  
Qy 2254 CCAATATCTCTTTGTTTAAACAAAAGATATTATTGTTGTATGAATCTAAATCAAGCC 2313  
Db 721 CCAATATCTCTTTGTTTAAACAAAAGATATTATTGTTGTATGAATCTAAATCAAGCC 780  
Qy 2314 CATCTGTCATTATGTTACTGCTTTTTTAATCATGTGGTTTTGTATATTAATTAATTTGTTG 2373  
Db 781 CATCTGTCATTATGTTACTGCTTTTTTAATCATGTGGTTTTGTATATTAATTAATTAATTTGTTG 840  
Qy 2374 ACTTCTTAGATTCACCTCCATATGTAATGTAAGCTCTTAATCTATGCTCTTTGTAATG 2433  
Db 841 ACTTCTTAGATTCACCTCCATATGTAATGTAAGCTCTTAATCTATGCTCTTTGTAATG 900  
Qy 2434 TGTAAATTTCTTCTGAAATAAAACCAATTTGTGAATAT 2470  
Db 901 TGTAAATTTCTTCTGAAATAAAACCAATTTGTGAATAT 937

RESULT 2  
US-11-266-748A-288390  
; Sequence 288390, Application US/11266748A  
; Publication No. US20060134663A1  
; GENERAL INFORMATION:  
; APPLICANT: Harkin, Paul  
; APPLICANT: Johnston, Patrick  
; APPLICANT: Mulligan, Karl  
; TITLE OF INVENTION: Transcriptome Microarray Technology and  
; TITLE OF INVENTION: Methods of Using the Same  
; FILE REFERENCE: 55815-0102 (319189)  
; CURRENT APPLICATION NUMBER: US/11/266.748A  
; CURRENT FILING DATE: 2005-11-03

; PRIOR APPLICATION NUMBER: EP 04105479.2  
; PRIOR FILING DATE: 2004-11-03  
; PRIOR APPLICATION NUMBER: EP 04105482.6  
; PRIOR FILING DATE: 2004-11-03  
; PRIOR APPLICATION NUMBER: EP 04105483.4  
; PRIOR FILING DATE: 2004-11-03  
; PRIOR APPLICATION NUMBER: EP 04105507.0  
; PRIOR FILING DATE: 2004-11-03  
; PRIOR APPLICATION NUMBER: EP 04105485.9  
; PRIOR FILING DATE: 2004-11-03  
; PRIOR APPLICATION NUMBER: EP 04105484.2  
; PRIOR FILING DATE: 2004-11-03  
; PRIOR APPLICATION NUMBER: US 60/662,276  
; PRIOR FILING DATE: 2005-03-14  
; PRIOR APPLICATION NUMBER: US 60/700,293  
; PRIOR FILING DATE: 2005-07-18  
; NUMBER OF SEQ ID NOS: 483996  
; SOFTWARE: PatentIn version 3.3  
; SEQ ID NO 288390  
; LENGTH: 1000  
; TYPE: DNA  
; ORGANISM: Homo Sapiens  
US-11-266-748A-288390

Query Match 37.9%; Score 937; DB 8; Length 1000;  
Best Local Similarity 100.0%; Pred. No. 2.6e-236;  
Matches 937; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1534 GAGAATACTACTAGCGCCAAATCGTTACACTACACCCCTCAAAGCTAGAAACCACTGCG 1593  
Db 1 GAGAATACTACTAGCGCCAAATCGTTACACTACACCCCTCAAAGCTAGAAACCACTGCG 60  
Qy 1594 TGAAGAAACTCCAATTTAAATACCAAGTAAATTTCAACAGGAAACAGACAAGTTAATGACAG 1653  
Db 61 TGAAGAAACTCCAATTTAAATACCAAGTAAATTTCAACAGGAAACAGACAAGTTAATGACAG 120  
Qy 1654 GTGTCAATAGCCCTGAGAGCGGTGCGCTCAGTGGAAATGGATCTCAACAGACATA 1713  
Db 121 GTGTCAATAGCCCTGAGAGCGGTGCGCTCAGTGGAAATGGATCTCAACAGACATA 180  
Qy 1714 TGAAGAGACTCCAAAGAAAGAGGAGCCAAAGTGTGTTGGAGCCCTTGAAGGGGGTTGG 1773  
Db 181 TGAAGAGACTCCAAAGAAAGAGGAGCCAAAGTGTGTTGGAGCCCTTGAAGGGGGTTGG 240  
Qy 1774 ATAAGGTATCACTGTGCTCACCAGAGCAAAAGAGGGTCTGCCAGAGAGCGGCCCA 1833  
Db 241 ATAAGGTATCACTGTGCTCACCAGAGCAAAAGAGGGTCTGCCAGAGAGCGGCCCA 300  
Qy 1834 GAAGACTAAAGCTTCACTATAATGTGACTACAACTAGATAGTGAATCCAGATCAACTGT 1893  
Db 301 GAAGACTAAAGCTTCACTATAATGTGACTACAACTAGATAGTGAATCCAGATCAACTGT 360  
Qy 1894 TGAATGAATAATGTCTATTCTTCCAAAGAGCATGTTGACTTTGTACAAAAGGGTTATA 1953  
Db 361 TGAATGAATAATGTCTATTCTTCCAAAGAGCATGTTGACTTTGTACAAAAGGGTTATA 420  
Qy 1954 CACTGAAGTGTCAAAACACAGTCAAGTCAAGTCAAGTGGGAAAGTGACAATGCAATTTGAATTAAG 2013  
Db 421 CACTGAAGTGTCAAAACACAGTCAAGTCAAGTGGGAAAGTGACAATGCAATTTGAATTAAG 480  
Qy 2014 TGTGCCAGCTTCAAAAACCCGATGTGTGGGTATCAGGAGGCGCGCTTAAGGGCGATG 2073  
Db 481 TGTGCCAGCTTCAAAAACCCGATGTGTGGGTATCAGGAGGCGCGCTTAAGGGCGATG 540  
Qy 2074 CTGGGGTTTCAAAAGATTAAGTGAAGACATCTCTAGCTGCAAGGTATAATTCATGG 2133  
Db 541 CTGGGGTTTCAAAAGATTAAGTGAAGACATCTCTAGCTGCAAGGTATAATTCATGG 600  
Qy 2134 ATTCTTCCATCTGCGCGATGAGTGTGGGTGTGATACAGCCTACATAAAGACTTTATGA 2193  
Db 601 ATTCTTCCATCTGCGCGATGAGTGTGGGTGTGATACAGCCTACATAAAGACTTTATGA 660  
Qy 2194 TCGCTTTGATTTTAAAGTTCATTTGGAACTACCAACTTGTGTTCTTAAAGAGCTATCTTAAGA 2253



; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105485.9
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105484.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: US 60/662,276
; PRIOR FILING DATE: 2005-03-14
; PRIOR APPLICATION NUMBER: US 60/700,293
; PRIOR FILING DATE: 2005-07-18
; NUMBER OF SEQ ID NOS: 483996
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 399255
; LENGTH: 1000
; TYPE: DNA
; ORGANISM: Homo Sapiens
US-11-266-748A-399255

Query Match 37.9%; Score 937; DB 8; Length 1000;
Best Local Similarity 100.0%; Pred. No. 2.6e-236;
Matches 937; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1534 GAGAAATACCTACAGCCCAATCGTTACACTACACCCCTCAAAAGCTAGAAACCAAGTGC 1593
DB 1 GAGAAATACCTACAGCCCAATCGTTACACTACACCCCTCAAAAGCTAGAAACCAAGTGC 60

QY 1594 TGAAGAAACTCCAATTAAATACCAAGTAAATCAACAGGAACAGACAAGTTAATGACAG 1653
DB 61 TGAAGAAACTCCAATTAAATACCAAGTAAATCAACAGGAACAGACAAGTTAATGACAG 120

QY 1654 GTGTCATTAGCCCTGAGAGCGGTGCGCTCAGTGGAAATGGATCTCAACCAAGACACATA 1713
DB 121 GTGTCATTAGCCCTGAGAGCGGTGCGCTCAGTGGAAATGGATCTCAACCAAGACACATA 180

QY 1714 TGAAGAGACTCAAAAGAAAGGAGCCCAAGTGTGGAGCCCTTGAAGGGGGTGG 1773
DB 181 TGAAGAGACTCAAAAGAAAGGAGCCCAAGTGTGGAGCCCTTGAAGGGGGTGG 240

QY 1774 ATAAGGTTATCACTGTCTCACCAGAGCAAAAGAGGGTCTGCGCAGAGCGGCCCA 1833
DB 241 ATAAGGTTATCACTGTCTCACCAGAGCAAAAGAGGGTCTGCGCAGAGCGGCCCA 300

QY 1834 GAAGACTAAAGCTTCACTATAATGTGACTACAACTAGATAGTGAATCCAGATCAACTGT 1893
DB 301 GAAGACTAAAGCTTCACTATAATGTGACTACAACTAGATAGTGAATCCAGATCAACTGT 360

QY 1894 TGAATGAATTAATGTCTATTCTTCCAAAGAGCATGTGACTTTGTACAAAAGGGTTATA 1953
DB 361 TGAATGAATTAATGTCTATTCTTCCAAAGAGCATGTGACTTTGTACAAAAGGGTTATA 420

QY 1954 CACTGAAGTGTCAAAACACAGTCAGATTTTGGGAAAGTGACAATGCCAATTTGAATTAGAAG 2013
DB 421 CACTGAAGTGTCAAAACACAGTCAGATTTTGGGAAAGTGACAATGCCAATTTGAATTAGAAG 480

QY 2014 TGTGCCAGCTTCAAAAACCCGATGTGGTGGGTATCAGAGGAGCAGCGGCTTAAAGGGGATG 2073
DB 481 TGTGCCAGCTTCAAAAACCCGATGTGGTGGGTATCAGAGGAGCAGCGGCTTAAAGGGGATG 540

QY 2074 CTTGGGTTTCAAAAGATTAGTGGAGACATCTCTATCTAGCTGCAAGGTATATTGATGG 2133
DB 541 CTTGGGTTTCAAAAGATTAGTGGAGACATCTCTATCTAGCTGCAAGGTATATTGATGG 600

QY 2134 ATTCTTCCATCTCCCGGATGAGTGTGGGTGTGATACAGCTTACATAAAGACTGTTATGA 2193
DB 601 ATTCTTCCATCTCCCGGATGAGTGTGGGTGTGATACAGCTTACATAAAGACTGTTATGA 660

QY 2194 TCGCTTTGATTTAAAGTTCAATTTGGAACACTACCAATCTGTTTCTTAAAGAGCTATCTTAAGA 2253
DB 661 TCGCTTTGATTTAAAGTTCAATTTGGAACACTACCAATCTGTTTCTTAAAGAGCTATCTTAAGA 720

QY 2254 CCAATATCTCTTTGTTTTTAAACAAAGATATTATTTTGTGTATGAATCTAAATCAAGCC 2313
DB 721 CCAATATCTCTTTGTTTTTAAACAAAGATATTATTTTGTGTATGAATCTAAATCAAGCC 780

RESULT 5
US-11-266-748A-470301/c
; Sequence 470301, Application US/11266748A
; Publication No. US20060134663A1
; GENERAL INFORMATION:
; APPLICANT: Harkin, Paul
; APPLICANT: Johnston, Patrick
; APPLICANT: Mulligan, Karl
; TITLE OF INVENTION: Transcriptome Microarray Technology and
; TITLE OF INVENTION: Methods of Using the Same
; FILE REFERENCE: 55815-0102 (319189)
; CURRENT APPLICATION NUMBER: US/11/266,748A
; CURRENT FILING DATE: 2005-11-03
; PRIOR APPLICATION NUMBER: EP 04105479.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105482.6
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105483.4
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105507.0
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105485.9
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105484.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: US 60/662,276
; PRIOR FILING DATE: 2005-03-14
; PRIOR APPLICATION NUMBER: US 60/700,293
; PRIOR FILING DATE: 2005-07-18
; NUMBER OF SEQ ID NOS: 483996
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 470301
; LENGTH: 1000
; TYPE: DNA
; ORGANISM: Homo Sapiens
US-11-266-748A-470301

Query Match 37.9%; Score 937; DB 8; Length 1000;
Best Local Similarity 100.0%; Pred. No. 2.6e-236;
Matches 937; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1534 GAGAAATACCTACAGCCCAATCGTTACACTACACCCCTCAAAAGCTAGAAACCAAGTGC 1593
DB 1000 GAGAAATACCTACAGCCCAATCGTTACACTACACCCCTCAAAAGCTAGAAACCAAGTGC 941

QY 1594 TGAAGAAACTCCAATTAAATACCAAGTAAATCAACAGGAACAGACAAGTTAATGACAG 1653
DB 940 TGAAGAAACTCCAATTAAATACCAAGTAAATCAACAGGAACAGACAAGTTAATGACAG 881

QY 1654 GTGTCATTAGCCCTGAGAGCGGTGCGCTCAGTGGAAATGGATCTCAACCAAGACACATA 1713
DB 880 GTGTCATTAGCCCTGAGAGCGGTGCGCTCAGTGGAAATGGATCTCAACCAAGACACATA 821

QY 1714 TGAAGAGACTCCAAAGAAAGGAGCCCAAGTGTGGAGCCCTTGAAGGGGGTGG 1773
DB 820 TGAAGAGACTCCAAAGAAAGGAGCCCAAGTGTGGAGCCCTTGAAGGGGGTGG 761

QY 1774 ATAAGGTTATCACTGTCTCACCAGAGCAAAAGAGGGTCTGCGCAGAGCGGCCCA 1833
DB 760 ATAAGGTTATCACTGTCTCACCAGAGCAAAAGAGGGTCTGCGCAGAGCGGCCCA 701

QY 2314 CATCTGTCATTAGTACTGTCTTTTAAATCATGTGGTTTGTATATTAATAAATGTTG 2373
DB 781 CATCTGTCATTAGTACTGTCTTTTAAATCATGTGGTTTGTATATTAATAAATGTTG 840

QY 2374 ACTTTCTTAGATTCACCTCCATATCTGAATGTAAGCTCTTAACATGTCTCTTTGTAATG 2433
DB 841 ACTTTCTTAGATTCACCTCCATATCTGAATGTAAGCTCTTAACATGTCTCTTTGTAATG 900

QY 2434 TGTAAATTTCTTCTGAAATAAAACCAATTTGTGAATAT 2470
DB 901 TGTAAATTTCTTCTGAAATAAAACCAATTTGTGAATAT 937

QY 1834 GAAGACTAAAGCTTCACTATAATGTGACTACAACTAGATTAGTGAATCCAGATCAACTGT 1893  
DB 700 GAAGACTAAAGCTTCACTATAATGTGACTACAACTAGATTAGTGAATCCAGATCAACTGT 641  
QY 1894 TGAATCAAAATGAATGTCTATTCTTCCAAAGAACGATGTGACTTTGTACAAAAGGGTTATA 1953  
DB 640 TGAATCAAAATGAATGTCTATTCTTCCAAAGAACGATGTGACTTTGTACAAAAGGGTTATA 591  
QY 1954 CACTGAAGTGTCAAAACACAGTCAAGATTTTGGGAAAGTGACAAATGCAATTTGAATTAGAAG 2013  
DB 580 CACTGAAGTGTCAAAACACAGTCAAGATTTTGGGAAAGTGACAAATGCAATTTGAATTAGAAG 521  
QY 2014 TGTGCCAGCTTCAAAAACCGGATGTGGTGTATCAGAGGCGAGCGCTTAAGGGCGATG 2073  
DB 520 TGTGCCAGCTTCAAAAACCGGATGTGGTGTATCAGAGGCGAGCGCTTAAGGGCGATG 461  
QY 2074 CTTGGGTTTACAAAAGATTAGTGGAAAGACATCTCTAGCTGCAAGGTATTAATTCATGG 2133  
DB 460 CTTGGGTTTACAAAAGATTAGTGGAAAGACATCTCTAGCTGCAAGGTATTAATTCATGG 401  
QY 2134 ATTCTTCCATCTCGCCGGATGAGTGTGGGTGTGATACAGCTACATAAAAGACTGTTATGA 2193  
DB 400 ATTCTTCCATCTCGCCGGATGAGTGTGGGTGTGATACAGCTACATAAAAGACTGTTATGA 341  
QY 2194 TGCCTTTGAATTTAAAGTTCATTGGAACATCCAACTTGTTCATAAGAGCTATCTTAAGA 2253  
DB 340 TGCCTTTGAATTTAAAGTTCATTGGAACATCCAACTTGTTCATAAGAGCTATCTTAAGA 281  
QY 2254 CCAATATCTCTTTGTTTTTAAACAAAGATATTTTGTGTATGAATCTTAATCAAGCC 2313  
DB 280 CCAATATCTCTTTGTTTTTAAACAAAGATATTTTGTGTATGAATCTTAATCAAGCC 221  
QY 2314 CATCTGCTATTATGTTACTGTCTTTTAAATCATGTGGTTTTGTATTAATAATTTGTTG 2373  
DB 220 CATCTGCTATTATGTTACTGTCTTTTAAATCATGTGGTTTTGTATTAATAATTTGTTG 161  
QY 2374 ACTTCTTAGATTCACCTCCATATGTAATGTAAGCTCTTAATGCTCTTTTGAATG 2433  
DB 160 ACTTCTTAGATTCACCTCCATATGTAATGTAAGCTCTTAATGCTCTTTTGAATG 101  
QY 2434 TGTAAATTTCTTCTGAAATTAACCAATTTGTGAATAT 2470  
DB 100 TGTAAATTTCTTCTGAAATTAACCAATTTGTGAATAT 64

## RESULT 6

US-11-266-748A-189434  
; Sequence 189434, Application US/11266748A  
; Publication No. US20060134663A1  
; GENERAL INFORMATION:  
; APPLICANT: Harkin, Paul  
; APPLICANT: Johnstone, Patrick  
; APPLICANT: Mulligan, Karl  
; TITLE OF INVENTION: Transcriptome Microarray Technology and  
; FILE REFERENCE: 55815-0102 (319189)  
; CURRENT APPLICATION NUMBER: US/11/266,748A  
; CURRENT FILING DATE: 2005-11-03  
; PRIOR APPLICATION NUMBER: EP 04105479.2  
; PRIOR FILING DATE: 2004-11-03  
; PRIOR APPLICATION NUMBER: EP 04105482.6  
; PRIOR FILING DATE: 2004-11-03  
; PRIOR APPLICATION NUMBER: EP 04105483.4  
; PRIOR FILING DATE: 2004-11-03  
; PRIOR APPLICATION NUMBER: EP 04105507.0  
; PRIOR FILING DATE: 2004-11-03  
; PRIOR APPLICATION NUMBER: EP 04105485.9  
; PRIOR FILING DATE: 2004-11-03  
; PRIOR APPLICATION NUMBER: EP 04105484.2  
; PRIOR FILING DATE: 2004-11-03  
; PRIOR APPLICATION NUMBER: US 60/662,276  
; PRIOR FILING DATE: 2005-03-14

; PRIOR APPLICATION NUMBER: US 60/700,293  
; PRIOR FILING DATE: 2005-07-18  
; NUMBER OF SEQ ID NOS: 48396  
; SOFTWARE: PatentIn version 3.3  
; SEQ ID NO 189434  
; LENGTH: 926  
; TYPE: DNA  
; ORGANISM: Homo Sapiens  
US-11-266-748A-189434

Query Match 36.5%; Score 901.6; DB 8; Length 926;  
Best Local Similarity 99.5%; Pred. No. 5.1e-227;  
Matches 915; Conservative 0; Mismatches 4; Indels 1; Gaps 1;  
QY 33 CGAAAAGATTCTTAGGAACGCCGTACAGCCGCGTCTCTCAGGACAGCAGGCGCCCTGTCC 92  
DB 5 CGAAAAGATTCTTAGGAACGCCGTACAGCCGCGTCTCTCAGGACAGCAGGCGCCCTGTCC 64  
QY 93 TTCTGTGGGCGCGCTCAGCCGCGTCCGCCCTCAGGTTCTTTTCTTAATTCAAA 152  
DB 65 TTCTGTGGGCGCGCTCAGCCGCGTCCGCCCTCAGGTTCTTTTCTTAATTCAAA 124  
QY 153 TAAACTTGCACAGGACTATGAAGATTATGATGAACCTTCTCAAAATATTATGAATTACAT 212  
DB 125 TAAACTTGCACAGGACTATGAAGATTATGATGAACCTTCTCAAAATATTATGAATTACAT 184  
QY 213 GAAACTATTGGACAGGCGCTTTGCAAGGTCAAACTTTGCCCTGCCATATCCTTACTGGA 272  
DB 185 GAAACTATTGGACAGGCGCTTTGCAAGGTCAAACTTTGCCCTGCCATATCCTTACTGGA 244  
QY 273 GAGATGTAGCTATAAAATCATGGATAAAACACACATAGGAGTGAATTTGCCCGGATC 332  
DB 245 GAGATGTAGCTATAAAATCATGGATAAAACACACATAGGAGTGAATTTGCCCGGATC 304  
QY 333 AAAACGAGATTGAGGCGCTTGAAGAACCTTGACACATATATGTCAACTCTTACCAT 392  
DB 305 AAAACGAGATTGAGGCGCTTGAAGAACCTTGACACATATATGTCAACTCTTACCAT 364  
QY 393 GTGCTAGACAGCAGCAACAAAATATTATGTTCTTGAGTACTGCCCTGGAGGAGCTG 452  
DB 365 GTGCTAGACAGCAGCAACAAAATATTATGTTCTTGAGTACTGCCCTGGAGGAGCTG 424  
QY 453 TTTGACTATATAATTTCCAGGATCGCTGTGACAGGAGGAGACCCGGGTTGTCTTCCT 512  
DB 425 TTTGACTATATAATTTCCAGGATCGCTGTGACAGGAGGAGACCCGGGTTGTCTTCCT 484  
QY 513 CAGATAGTATCTGTGTTGCTTATGTCACAGCAGCGGCTATGCTCACAGGACCTCAAG 572  
DB 485 CAGATAGTATCTGTGTTGCTTATGTCACAGCAGCGGCTATGCTCACAGGACCTCAAG 544  
QY 573 CCAGAAAATTTGCTGTTGATGAATATATAAATTTAAAGCTGATTTGGTCTCTGT 632  
DB 545 CCAGAAAATTTGCTGTTGATGAATATATAAATTTAAAGCTGATTTGGTCTCTGT 604  
QY 633 GCAAAACCAAGGTAACAAGGATTAACATCTACAGACATGCTGTGGGAGTCTGGCTTAT 692  
DB 605 GCAAAACCAAGGTAACAAGGATTAACATCTACAGACATGCTGTGGGAGTCTGGCTTAT 664  
QY 693 GCAGCACCTGAGTTAATACAGGCAATCATATCTTGGATCAGAGGAGATGTTGGAGC 752  
DB 665 GCAGCACCTGAGTTAATACAGGCAATCATATCTTGGATCAGAGGAGATGTTGGAGC 724  
QY 753 ATGGGCATATCTGTTATATGTTCTTATGTTGATTTTACACATTTTGAATGATATATGA 812  
DB 725 ATGGGCATATCTGTTATATGTTCTTATGTTGATTTTCTACCATTTTGAATGATATATGA 784  
QY 813 ATGGCTTTTATCAAGAGATTATGAGAGAAAATATGATGTTTCCAGTGGCTCTCTCCC 872  
DB 785 ATGGCTTTTATCAAGAGATTATGAGAGAAAATATGATGTTTCCAGTGGCTCTCTCCC 844  
QY 873 AGTAGCATTTCTCTTCTTCAAAATGCTGAGGTGGAGCCCAAGAAACGATTTCTATG 932  
DB 845 AGTAGCATTTCTCTTCTTCAAAATGCTGCA-GTGGACCCCAAGAAACGATTTCTATG 903

QY 933 AAAAAATCTATTGAACCATCC 952  
Db 904 AAAAAATCTATTGAACCATCC 923

RESULT 7  
US-11-266-748A-194943  
; Sequence 194943, Application US/11266748A  
; Publication No. US20060134663A1  
; GENERAL INFORMATION:  
; APPLICANT: Harkin, Paul  
; APPLICANT: Johnston, Patrick  
; APPLICANT: Mulligan, Karl  
; TITLE OF INVENTION: Transcriptome Microarray Technology and  
; TITLE OF INVENTION: Methods of Using the Same  
; FILE REFERENCE: 55815-0102 (319189)  
; CURRENT APPLICATION NUMBER: US/11/266, 748A  
; CURRENT FILING DATE: 2005-11-03  
; PRIOR APPLICATION NUMBER: EP 04105479.2  
; PRIOR FILING DATE: 2004-11-03  
; PRIOR APPLICATION NUMBER: EP 04105482.6  
; PRIOR FILING DATE: 2004-11-03  
; PRIOR APPLICATION NUMBER: EP 04105483.4  
; PRIOR FILING DATE: 2004-11-03  
; PRIOR APPLICATION NUMBER: EP 04105507.0  
; PRIOR FILING DATE: 2004-11-03  
; PRIOR APPLICATION NUMBER: EP 04105485.9  
; PRIOR FILING DATE: 2004-11-03  
; PRIOR APPLICATION NUMBER: EP 04105484.2  
; PRIOR FILING DATE: 2004-11-03  
; PRIOR APPLICATION NUMBER: US 60/662, 276  
; PRIOR FILING DATE: 2005-03-14  
; PRIOR APPLICATION NUMBER: US 60/700, 293  
; NUMBER OF SEQ ID NOS: 483996  
; SOFTWARE: PatentIn version 3.3  
; SEQ ID NO 194943  
; LENGTH: 926  
; TYPE: DNA  
; ORGANISM: Homo Sapiens  
US-11-266-748A-194943

Query Match 36.5%; Score 901.6; DB 8; Length 926;  
Best Local Similarity 99.5%; Pred. No. 5.1e-227;  
Matches 915; Conservative 0; Mismatches 4; Indels 1; Gaps 1;

QY 33 CGAAAAAGATTCTTAGGAACGCCGTACCCAGCGCGTCTCTCAGGACAGCAGGCGCCCTGTCC 92  
Db 5 CGAAAAAGATTCTTAGGAACGCCGTACCCAGCGCGTCTCTCAGGACAGCAGGCGCCCTGTCC 64

QY 93 TTCTGTGGCGCGCTCAGCGTGGCTCGCCCTCGCCCTCAGGTTCTTTTCTAATTCGAA 152  
Db 65 TTCTGTGGCGCGCTCAGCGTGGCTCGCCCTCGCCCTCAGGTTCTTTTCTAATTCGAA 124

QY 153 TAACTTGCAGAGGACTGAAGAATATGATGAACCTCTCAATATATGATTAATACAT 212  
Db 125 TAACTTGCAGAGGACTGAAGAATATGATGAACCTCTCAATATATGATTAATACAT 184

QY 213 GAAACTATTGGGACAGGTGGCTTTGCAAAAGGTCAAACTTGCCTGCCATATCCTTACTGGA 272  
Db 185 GAAACTATTGGGACAGGTGGCTTTGCAAAAGGTCAAACTTGCCTGCCATATCCTTACTGGA 244

QY 273 GAGATGGTAGCTATAAAATCATGGATAAAACACACTAGGAGTGAATTTGCCCGGATC 332  
Db 245 GAGATGGTAGCTATAAAATCATGGATAAAACACACTAGGAGTGAATTTGCCCGGATC 304

QY 333 AAAACGAGATTGAGGCTTTGAAGAACCTCAGACATCAGCATATATGTCAACTCTACCAT 392  
Db 305 AAAACGAGATTGAGGCTTTGAAGAACCTCAGACATCAGCATATATGTCAACTCTACCAT 364

QY 393 GTGCTAGAGACAGCCAAACAAAATATTTCATGGTTCTTGAGTACTGCCCTGGAGGAGCTG 452

Db 365 GTGCTAGAGACAGCCAAACAAAATATTTCATGGTTCTTGAGTACTGCCCTGGAGGAGCTG 424  
QY 453 TTTGACTATATAATTTCCAGGATCGCTGTGTCAGAAAGAGGAGACCCGGGTTCCTTCCCT 512  
Db 425 TTTGACTATATAATTTCCAGGATCGCTGTGTCAGAAAGAGGAGACCCGGGTTCCTTCCCT 484  
QY 513 CAGATAGTATCTGCTGTTGCTTATGTCACAGCCAGGCGTATGTCACAGGACCTCAAG 572  
Db 485 CAGATAGTATCTGCTGTTGCTTATGTCACAGCCAGGCGTATGTCACAGGACCTCAAG 544  
QY 573 CCAGAAAATTTGCTGTTGATGAATATCATATAATTAAGCTGATTTGCTTCTCTCT 632  
Db 545 CCAGAAAATTTGCTGTTGATGAATATCATATAATTAAGCTGATTTGCTTCTCTCT 604  
QY 633 GCAAAACCCAGGGTAAACAAGGATTACCATCTACAGACATGCTGTGGGAGTCTGGCTTAT 692  
Db 605 GCAAAACCCAGGGTAAACAAGGATTACCATCTACAGACATGCTGTGGGAGTCTGGCTTAT 664  
QY 693 GCAGCACCTGAGTTAATACAGGCAAAATCATATCTTTGGATCAGAGGACAGATGTTGGAGC 752  
Db 665 GCAGCACCTGAGTTAATACAGGCAAAATCATATCTTTGGATCAGAGGACAGATGTTGGAGC 724  
QY 753 ATGGGCATCTGTTATATGTTCTTATGTTGGATTTCTACCATTTGATGATGATTAATGA 812  
Db 725 ATGGGCATCTGTTATATGTTCTTATGTTGGATTTCTACCATTTGATGATGATTAATGA 784  
QY 813 ATGGCTTTTATACAGAAGATTATGAGAGGAAATATGATGTTCCCAAGTGGCTCTCTCCC 872  
Db 785 ATGGCTTTTATACAGAAGATTATGAGAGGAAATATGATGTTCCCAAGTGGCTCTCTCCC 844  
QY 873 AGTAGCATTTCTCTTCTTCAACAAATGTCAGGTGGAGCCCAAGAAACGGAATTTCTATG 932  
Db 845 AGTAGCATTTCTCTTCTTCAACAAATGTCGCA-GTGGACCCCAAGAAACGGAATTTCTATG 903

QY 933 AAAAAATCTATTGAACCATCC 952  
Db 904 AAAAAATCTATTGAACCATCC 923

RESULT 8  
US-11-266-748A-262797  
; Sequence 262797, Application US/11266748A  
; Publication No. US20060134663A1  
; GENERAL INFORMATION:  
; APPLICANT: Harkin, Paul  
; APPLICANT: Johnston, Patrick  
; APPLICANT: Mulligan, Karl  
; TITLE OF INVENTION: Transcriptome Microarray Technology and  
; TITLE OF INVENTION: Methods of Using the Same  
; FILE REFERENCE: 55815-0102 (319189)  
; CURRENT APPLICATION NUMBER: US/11/266, 748A  
; CURRENT FILING DATE: 2005-11-03  
; PRIOR APPLICATION NUMBER: EP 04105479.2  
; PRIOR FILING DATE: 2004-11-03  
; PRIOR APPLICATION NUMBER: EP 04105482.6  
; PRIOR FILING DATE: 2004-11-03  
; PRIOR APPLICATION NUMBER: EP 04105483.4  
; PRIOR FILING DATE: 2004-11-03  
; PRIOR APPLICATION NUMBER: EP 04105507.0  
; PRIOR FILING DATE: 2004-11-03  
; PRIOR APPLICATION NUMBER: EP 04105485.9  
; PRIOR FILING DATE: 2004-11-03  
; PRIOR APPLICATION NUMBER: EP 04105484.2  
; PRIOR FILING DATE: 2004-11-03  
; PRIOR APPLICATION NUMBER: US 60/662, 276  
; PRIOR FILING DATE: 2005-03-14  
; PRIOR APPLICATION NUMBER: US 60/700, 293  
; NUMBER OF SEQ ID NOS: 483996  
; SOFTWARE: PatentIn version 3.3  
; SEQ ID NO 262797  
; LENGTH: 1050  
; TYPE: DNA

; ORGANISM: Homo Sapiens		Db		920 GGGGCTCACCAGGAGCAAAAGAGGGTTCTGCCAAA 957	
; FEATURE:		RESULT 9			
; NAME/KEY: misc feature		US-11-266-748A-323314/c			
; LOCATION: (980)..(980)		; Sequence 323314, Application US/11266748A			
; OTHER INFORMATION: n is a, c, g, or t		; Publication No. US20060134663A1			
US-11-266-748A-262797		; GENERAL INFORMATION:			
		; APPLICANT: Harkin, Paul			
		; APPLICANT: Johnston, Patrick			
		; APPLICANT: Mulligan, Karl			
		; TITLE OF INVENTION: Transcriptome Microarray Technology and			
		; TITLE OF INVENTION: Methods of Using the Same			
		; FILE REFERENCE: 55815-0102 (319189)			
		; CURRENT APPLICATION NUMBER: US/11/266,748A			
		; CURRENT FILING DATE: 2005-11-03			
		; PRIOR APPLICATION NUMBER: EP 04105479.2			
		; PRIOR FILING DATE: 2004-11-03			
		; PRIOR APPLICATION NUMBER: EP 04105482.6			
		; PRIOR FILING DATE: 2004-11-03			
		; PRIOR APPLICATION NUMBER: EP 04105483.4			
		; PRIOR FILING DATE: 2004-11-03			
		; PRIOR APPLICATION NUMBER: US 60/662,276			
		; PRIOR FILING DATE: 2005-03-14			
		; PRIOR APPLICATION NUMBER: US 60/700,293			
		; PRIOR FILING DATE: 2005-07-18			
		; NUMBER OF SEQ ID NOS: 483996			
		; SOFTWARE: PatentIn version 3.3			
		; SEQ ID NO 323314			
		; LENGTH: 1050			
		; TYPE: DNA			
		; ORGANISM: Homo Sapiens			
		; FEATURE:			
		; NAME/KEY: misc feature			
		; LOCATION: (71)..(71)			
		; OTHER INFORMATION: n is a, c, g, or t			
		US-11-266-748A-323314			
		Query Match		32.0%; Score 791.2; DB 8; Length 1050;	
		Best Local Similarity		97.2%; Pred. No. 6.3e-198;	
		Matches 912; Conservative		0; Mismatches 13; Indels 13; Gaps 10;	
		Qy		899 GCTGAGGTGGACCCAAAGAACCGATTTCTATGAAAAATCTATTGAAACCATCCCTGGAT 958	
		Db		1031 GCACGAGGTGGACCCAAAGAACCGATTTCTATGAAAAATCTATTGAAACCATCCCTGGAT 972	
		Qy		959 CATGCAAGATTACAACTATCTCTGTTGAGTGGCAAGCAAGATCTCTTTTATTTCACCTCGA 1018	
		Db		971 CATGCAAGATTACAACTATCTCTGTTGAGTGGCAAGCAAGATCTCTTTTATTTCACCTCGA 912	
		Qy		1019 TGATGATTGCGTAAACAGAACTTTCTGTATCATCACAGAAACCAACAGGCAAAACAATGGAGGA 1078	
		Db		911 TGATGATTGCGTAAACAGAACTTTCTGTATCATCACAGAAACCAACAGGCAAAACAATGGAGGA 852	
		Qy		1079 TTTAATTTTCACTGTGGCAGTATGATCAGCTACGGCTACCTATCTTCTGCTTCTTAGCCAA 1138	
		Db		851 TTTAATTTTCACTGTGGCAGTATGATCAGCTACGGCTACCTATCTTCTGCTTCTTAGCCAA 792	
		Qy		1139 GAAGGCTCGGGGAAACCAAGTTGGTTTAAAGGCTTCTTTCTCTCTCTGGACAGCCAG 1198	
		Db		791 GAAGGCTCGGGGAAACCAAGTTGGTTTAAAGGCTTCTTTCTCTCTCTGGACAGCCAG 732	
		Qy		1199 TGCTACCCCATTCACAGACATCAAGTCAAATAATTGGAGTCTGGAAAGATGTGACCCGCAAG 1258	
		Db		731 TGCTACCCCATTCACAGACATCAAGTCAAATAATTGGAGTCTGGAAAGATGTGACCCGCAAG 672	
		Qy		1259 TGATAAAATTTATGTGGCGGGATTAAATAGATATGATTTGGTGTGAAGATGATTTATCAAC 1318	



```
Db 671 TGATAAAATATGTGGCGGATTAATAGACTATGATGGTGGAAGATGATTTATCAAC 612
Qy 1319 AGGTGCTGCTACTCCCGAACATCAGATTGTACCAAGTACTGACAGAGATCAATCAATGGGCT 1378
Db 611 AGGTGCTGCTACTCCCGAACATCAGATTGTACCAAGTACTGACAGAGATCAATCAATGGGCT 552
Qy 1379 GGAATCTAAATCAATTAACCTCCAGCCTTATGACAGACACCTGCAAAATAAATTAAGAAACA 1438
Db 551 GGAATCTAAATCAATTAACCTCCAGCCTTATGACAGACACCTGCAAAATAAATTAAGAAACA 492
Qy 1439 AGAAATGTATATACCTCTTAAGTCTGCTGTAAGAAATGAAGAGTACTTTATGTTTCTCTGA 1498
Db 491 AGAAATGTATATACCTCTTAAGTCTGCTGTAAGAAATGAAGAGTACTTTATGTTTCTCTGA 432
Qy 1499 GCCAAGAGCTCCAGTTAATGAACACAGCATTAAGAGAGAAATACTCACTACGCCAATCG 1558
Db 431 GCCAAGAGCTCCAGTTAATGAACACAGCATTAAGAGAGAAATACTCACTACGCCAATCG 372
Qy 1559 TTACACTACACCTCAAAAGCTAGAACACAGTGCCTGAAAGAACTCCCAATTAATAATACC 1618
Db 371 TTACACTACACCTCAAAAGCTAGAACACAGTGCCTGAAAGAACTCCCAATTAATAATACC 312
Qy 1619 AGTAAATTTCAACA-GGAACAGACAAGTTAATGAAGAGTGTCTATTAG-CCCTGAGAGCGG 1676
Db 311 AGTAAATTTCAACAGGAGCAAGCAAGTTAATGAAGAGTGTCTATTAG-CCCTGAGAGCGG 252
Qy 1677 TG-CCGCTCAGTGGA--TTGGATCTCAACCAAGACATAT-GGAGAGACTCCAAAAG 1732
Db 251 TGCCCGCTCAGTGGAATTTGGGATCTCAACCAAGACATATGGGAGGAGACTCCAAAAG 192
Qy 1733 AAA-GGGAGCCAAAGTGT-TT-GGGAGCTTTGAAAGGGGGTTGG--ATAAGGTTTACT 1787
Db 191 AAAGGGAGCCAAAGTGT-TT-GGGAGCTTTGAAAGGGGGTTGGAATAGGGTTTACT 132
Qy 1788 -GTGCTCACCAGGAGC-AAAAGGAAGGGTTCTCCACGA 1823
Db 131 GGGGCTCACCAGGAGCAAAAGGAAGGGTTCTGCCAAA 94
```

```
RESULT 10
US-11-266-748A-9105
; Sequence 9105, Application US/11266748A
; Publication No. US20060134663A1
; GENERAL INFORMATION:
; APPLICANT: Harkin, Paul
; APPLICANT: Johnston, Patrick
; APPLICANT: Mulligan, Karl
; TITLE OF INVENTION: Transcriptome Microarray Technology and
; TITLE OF INVENTION: Methods of Using the Same
; FILE REFERENCE: 5815-0102 (319189)
; CURRENT APPLICATION NUMBER: US/11/266, 748A
; CURRENT FILING DATE: 2005-11-03
; PRIOR APPLICATION NUMBER: EP 04105479.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105482.6
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105483.4
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105507.0
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105485.9
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105484.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: US 60/662,276
; PRIOR FILING DATE: 2005-03-14
; PRIOR APPLICATION NUMBER: US 60/700,293
; PRIOR FILING DATE: 2005-07-18
; NUMBER OF SEQ ID NOS: 48396
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 9105
; LENGTH: 758
```

```
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (427)..(427)
; OTHER INFORMATION: n is a, c, g, or t
US-11-266-748A-9105

Query Match 29.1%; Score 717.6; DB 8; Length 758;
Best Local Similarity 98.5%; Pred. No. 1.3e-178;
Matches 745; Conservative 0; Mismatches 9; Indels 2; Gaps 2;

Qy 1715 GGAGGAGACTCCAAAAGAAAGAGGAGCCAAAGTGTGTGGAGCCTTGAAGGGGGTTGA 1774
Db 3 GGAGGAGACTCCAAAAGAAAGAGGAGCCAAAGTGTGTGGAGCCTTGAAGGGGGTTGA 62
Qy 1775 TAAAGTTATCACTGTGCTCACCAGAGCAAAAGGAAGGGTTCTGCAGAGACGGGCCAG 1834
Db 63 TAAAGTTATCACTGTGCTCACCAGAGCAAAAGGAAGGGTTCTGCAGAGACGGGCCAG 122
Qy 1835 AAGACTTAAAGCTTCACTATAATGTGACTACAACTAGATTAGTGAATCCAGATCAACTGTT 1894
Db 123 AAGACTTAAAGCTTCACTATAATGTGACTACAACTAGATTAGTGAATCCAGATCAACTGTT 182
Qy 1895 GAATGAAATAATGTCTATTCTTCCAAAGAGCATGTGATTTGTACAAAAGGGTTATAC 1954
Db 183 GAATGAAATAATGTCTATTCTTCCAAAGAGCATGTGATTTGTACAAAAGGGTTATAC 242
Qy 1955 ACTGAAGTGTCAAAACACAGTCAAGTTTGGGAAAGTGACAAATTTGAATTTAGAAGT 2014
Db 243 ACTGAAGTGTCAAAACACAGTCAAGTTTGGGAAAGTGACAAATTTGAATTTAGAAGT 302
Qy 2015 GTGCCAGCTTCAAAACCCGATGTGGTGTATCAGGAGGAGCGGCTTAAAGGGCATGC 2074
Db 303 GTGCCAGCTTCAAAACCCGATGTGGTGTATCAGGAGGAGCGGCTTAAAGGGCATGC 362
Qy 2075 CTGGGTTTACAAAAGATTAGTGGAAAGACATCTATCTAGCTCAAGGTATAATTTGATGA 2134
Db 363 CTGGGTTTACAAAAGATTAGTGGAAAGACATCTATCTAGCTCAAGGTATAATTTGATGA 422
Qy 2135 TTCT-TCCATCTCGCGGATGAGTGGTGATGATACAGCCTACATAAGAGCTTTATGA 2193
Db 423 TTCTTCCATCTCGCGGATGAGTGGTGATGATACAGCCTACATAAGAGCTTTATGA 482
Qy 2194 TCGCTTTGATTTTAAAGTTCAATTGGAACTACCAACTTTGTTCTTAAAGAGCTATCTTAAGA 2253
Db 483 TCGCTTTGATTTTAAAGTTCAATTGGAACTACCAACTTTGTTCTTAAAGAGCTATCTTAAGA 542
Qy 2254 CCAATATCTCTTTGTTTAAACAAAAGATATTTTGTGTATGAATCTAAATCAAGCC 2313
Db 543 CCAATATCTCTTTGTTTAAACAAAAGATATTTTGTGTATGAATCTAAATCAAGCC 602
Qy 2314 CATCTGTCAATATGTTACTGCTTTTAAATCATGGGGTTTGGATATTAAATTTGGTG 2373
Db 603 CATCTGTCAATATGTTACTGCTTTTAAATCATGGGGTTTGGATATTAAATTTGGTG 662
Qy 2374 ACTTCTTAGATTCTCTCCATATGTAATGTAAGTCTTTAACT-ATGCTCTTCTTGAAT 2432
Db 663 ACTTCTTAGATTCTCTCCATATGTAATGTAAGTCTTTAACTTAACTATGCTCTTTCGAAT 722
Qy 2433 GTGTAATTTCTTCTGAAATAAAACCAATTTGTAAT 2468
Db 723 GTGTAATTTCTCTGAAATAAAACCAATTTGTGCAT 758

RESULT 11
US-11-266-748A-189435
; Sequence 189435, Application US/11266748A
; Publication No. US20060134663A1
; GENERAL INFORMATION:
; APPLICANT: Harkin, Paul
; APPLICANT: Johnston, Patrick
; APPLICANT: Mulligan, Karl
```

```

; TITLE OF INVENTION: Transcriptome Microarray Technology and
; TITLE OF INVENTION: Methods of Using the Same
; FILE REFERENCE: 55815-0102 (319189)
; CURRENT APPLICATION NUMBER: US/11/266,748A
; CURRENT FILING DATE: 2005-11-03
; PRIOR APPLICATION NUMBER: EP 04105479.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105482.6
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105483.4
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105507.0
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105485.9
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105484.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: US 60/662,276
; PRIOR FILING DATE: 2005-03-14
; PRIOR APPLICATION NUMBER: US 60/700,293
; PRIOR FILING DATE: 2005-07-18
; NUMBER OF SEQ ID NOS: 48396
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 189435
; LENGTH: 948
; TYPE: DNA
; ORGANISM: Homo Sapiens
US-11-266-748A-189435

```

Query Match	27.7%	Score 685.4	DB 8	Length 948
Best Local Similarity	99.0%	Pred. No. 4.3e-170		
Matches 722	Conservative 0	Mismatches 1	Indels 6	Gaps 3
Qy	576	GAATAATTTGCTGTTTGATGAATATCATATAATTAAGCTGATTGACTTTGGTCTCTGTGCA	635	
Db	222	GAAATTTGCTGTTTGATGAATATCATATAATTAAGCTGATTGACTTTGGTCTCTGTGCA	281	
Qy	636	AAACCCCA----AGGGTAACAAGGATTACCATCTACAGACATGCTGTGGAGTCTGGCTTA	691	
Db	282	AAACCCCAAGCGAGGTAACAAGGATTACCATCTACAGACATGCTGTGGAGTCTGGCTTA	341	
Qy	692	TGCAGCACCTGAGTTAATACAAAGCAATCATATCTTGGATCAGAGCAGATGTTGGAG	751	
Db	342	TGCAGCACCTGAGTTAATACAAAGCAATCATATCTTGGATCAGAGCAGATGTTGGAG	401	
Qy	752	CATGGGCATACTGTTATATGTTCTTATGTGTGGATTTCTACCATTTGATGATGATAATGT	811	
Db	402	CATGGGCATACTGTTATATGTTCTTATGTGTGGATTTCTACCATTTGATGATGATAATGT	461	
Qy	812	AATGGCTTTTATACAAGAAGATTATGAGAGGAAAAATATGATGTTCCCAAGTGGCTCTCTCC	871	
Db	462	AATGGCTTTTATACAAGAAGATTATGAGAGGAAAAATATGATGTTCCCAAGTGGCTCTCTCC	521	
Qy	872	CAGTAGCATTTCTGCTTCTTCAACAAATGCTGCAGGTGGAACCAAGAAACGGATTTCTAT	931	
Db	522	CAGTAGCATTTCTGCTTCTTCAACAAATGCTGCAGGTGGAACCAAGAAACGGATTTCTAT	581	
Qy	932	GAATAATCTATTGAACCATCCCTGGATCATGCAAGATTACAACTATCTCTGTGTAGTGCA	991	
Db	582	GAATAATCTATTGAACCATCCCTGGATCATGCAAGATTACAACTATCTCTGTGTAGTGCA	641	
Qy	992	AAGCAAGAAATCCCTTTTATTCACCTCGATGATGATTGGGTAAACGAACTTCTGTACATCA	1051	
Db	642	ARGCAAGAAATCCCTTTTATTCACCTCGATGATGATTGGGTAAACGAACTTCTGTACATCA	701	
Qy	1052	CAGAAACCAACAGGCAAAACAATGGAGGATTTAAATTTCACTGTGGCAGTATGATCACTCAC	1111	
Db	702	CAGAAACCAACAGGCAAAACAATGGAGGATTTAAATTTCACTGTGGCAGTATGATCACTCAC	761	
Qy	1112	GGCTACTTATCTTCTGCTTCTAGCCAAAGAGGCTCGGGGAAAAACAGTTCGTTTAAAGCT	1171	
Db	762	GGCTACTTATCTTCTGCTTCTAGCCAAAGAA-GCTCGGGGAAAAACAGTTCGTTTAAAGCT	820	

Qy	1172	TTCTTTCTTTCTCTGTGGACAAGCCAGTGTCTACCCATTACAGACATCAAGTCAATAAA	1231
Db	821	TTCTTTCTTTCTCTGTGGACAAGCCAGTGTCTACCCATTACAGACATCAAGTCAATAAA	880
Qy	1232	TTGAGTCTCGGAGAGTGTGACCCCAAGTGTATAAAATTTATGTGCGGGATTATAGACTA	1291
Db	881	TTGAGTCTCGGAGAGTGTGACCCCAAGTGTATAAAATTTATGTGCGGGATT-ATAGACTA	939
Qy	1292	TGATTGGTG 1300	
Db	940	TGATTGGTG 948	

```

RESULT 12
US-11-266-748A-194944
; Sequence 194944, Application US/11266748A
; Publication No. US20060134663A1
; GENERAL INFORMATION:
; APPLICANT: Harkin, Paul
; APPLICANT: Johnston, Patrick
; APPLICANT: Mulligan, Karl
; TITLE OF INVENTION: Transcriptome Microarray Technology and
; TITLE OF INVENTION: Methods of Using the Same
; FILE REFERENCE: 5815-0102 (319189)
; CURRENT APPLICATION NUMBER: US/11/266,748A
; CURRENT FILING DATE: 2005-11-03
; PRIOR APPLICATION NUMBER: EP 04105479.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105482.6
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105483.4
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105507.0
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105485.9
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105484.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: US 60/662,276
; PRIOR FILING DATE: 2005-03-14
; PRIOR APPLICATION NUMBER: US 60/700,293
; PRIOR FILING DATE: 2005-07-18
; NUMBER OF SEQ ID NOS: 483996
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 194944
; LENGTH: 948
; TYPE: DNA
; ORGANISM: Homo Sapiens
US-11-266-748A-194944

```

Query Match	27.7%;	Score	685.4;	DB	8;	Length	948;
Best Local Similarity	99.0%;	Pred. No.	4.3e-170;				
Matches	722;	Conservative	0;	Mismatches	1;	Indels	6;
Gaps	3;						

Qy	576	GAATAATTGCTGTTTGATCAATATCATATAATTAAGCTGATTGACATTGGTCTCTCTGTGCA	635
Db	222	GAATAATTGCTGTTTGATGAATATCATATAATTAAGCTGATTGACATTGGTCTCTGTGCA	281
Qy	636	AAACCCA-----AGGGTAAACAAGATTACCATCTACAGACATGTGTGGGAGTCTGGCTTA	691
Db	282	AAACCCAAGGCAAGGGTAAACAAGATTACCATCTACAGACATGTGTGGGAGTCTGGCTTA	341
Qy	692	TGCAGCACCTGAGTTAATACAGGCAANTCATATCTTGGATCAGAGGCAGATGTTTGGAG	751
Db	342	TGCAGCACCTGAGTTAATACAGGCAANTCATATCTTGGATCAGAGGCAGATGTTTGGAG	401
Qy	752	CATGGGCATACCTGTTATATGTTCTTATGTGTGGATTCTACCAATTTGATGATGATAATGT	811
Db	402	CATGGGCATACCTGTTATATGTTCTTATGTGTGGATTCTACCAATTTGATGATGATAATGT	461
Qy	812	AATGGCTTTTATACAAGAAGATTATGAGAGAAATATGATGTTCCCAAGTGGCTCTCTCC	871
Db	462	AATGGCTTTTATACAAGAAGATTATGAGAGAAATATGATGTTCCCAAGTGGCTCTCTCC	521

		Matches	696;	Conservative	0;	Mismatches	14;	Indels	87;	Gaps	1;
QY	872	CAGTAGCATTCCTCTTCTTCAACAAATGCTGCAGGTGGACCCAAAGAAACGGATTCTAT									
Db	522	CAGTAGCATTCCTCTTCTTCAACAAATGCTGCAGGTGGACCCAAAGAAACGGATTCTAT									
QY	932	GAAAAATCTATTGAACCATCCCTGGATCATGCAAGATTACAATATCTCTGTTGAGTGGA									
Db	582	GAAAAATCTATTGAACCATCCCTGGATCATGCAAGATTACAATATCTCTGTTGAGTGGA									
QY	992	AGCAAGAATCCTTTTATTCACCTCATGATGATGCGTAAACAGAACTTTCTGTACATCA									
Db	642	AGCAAGAATCCTTTTATTCACCTCATGATGATGCGTAAACAGAACTTTCTGTACATCA									
QY	1052	CAGAAACAACAGGCAACCAATGAGGATTTAAATTCACCTGTGGCAGTATGATCACCCTCAC									
Db	702	CAGAAACAACAGGCAACCAATGAGGATTTAAATTCACCTGTGGCAGTATGATCACCCTCAC									
QY	1112	GGCTACCTATCTCTCTCTTCTAGCCCAAGAGGCTCGGGGAAACACCAAGTTGCTTTAAGGCT									
Db	762	GGCTACCTATCTCTCTCTTCTAGCCCAAGAA-GCTCGGGGAAACACCAAGTTGCTTTAAGCT									
QY	1172	TTCTTCTTCTCTGTGGCAAGCCAGTGTACCCCAATTCACAGACATCAAGTCAAAATTA									
Db	821	TTCTTCTTCTCTGTGGCAAGCCAGTGTACCCCAATTCACAGACATCAAGTCAAAATTA									
QY	1232	TTGGAGTCTGGAGATGTGACCCCAAGTGATTAATAAATTTATGTGGCGGATTAATAGACTA									
Db	881	TTGGAGTCTGGAGATGTGACCCCAAGTGATTAATAAATTTATGTGGCGGATTAATAGACTA									
QY	1292	TGATTGGTG 1300									
Db	940	TGATTGGTG 948									

RESULT 14  
US-11-266-748A-44240  
; Sequence 44240, Application US/11266748A  
; Publication No. US20060134663A1  
; GENERAL INFORMATION:  
; APPLICANT: Harkin, Paul  
; APPLICANT: Johnston, Patrick  
; APPLICANT: Mulligan, Karl  
; TITLE OF INVENTION: Transcriptome Microarray Technology and  
; TITLE OF INVENTION: Methods of Using the Same  
; FILE REFERENCE: 55815-0102 (319189)  
; CURRENT APPLICATION NUMBER: US/11/266, 748A  
; PRIOR FILING DATE: 2005-11-03  
; PRIOR APPLICATION NUMBER: EP 04105479.2

Query Match 23.9%; Score 590.6; DB 8; Length 728;  
Best Local Similarity 87.3%; Pred. No. 3.4e-145;

;; PRIOR FILING DATE: 2004-11-03  
;; PRIOR APPLICATION NUMBER: EP 04105482.6  
;; PRIOR FILING DATE: 2004-11-03  
;; PRIOR APPLICATION NUMBER: EP 04105483.4  
;; PRIOR FILING DATE: 2004-11-03  
;; PRIOR APPLICATION NUMBER: EP 04105507.0  
;; PRIOR FILING DATE: 2004-11-03  
;; PRIOR APPLICATION NUMBER: EP 04105485.9  
;; PRIOR FILING DATE: 2004-11-03  
;; PRIOR APPLICATION NUMBER: EP 04105484.2  
;; PRIOR FILING DATE: 2004-11-03  
;; PRIOR APPLICATION NUMBER: US 60/662,276  
;; PRIOR FILING DATE: 2005-03-14  
;; PRIOR APPLICATION NUMBER: US 60/700,293  
;; PRIOR FILING DATE: 2005-07-18  
;; NUMBER OF SEQ ID NOS: 48396  
;; SOFTWARE: PatentIn version 3.3  
;; SEQ ID NO 44240  
;; LENGTH: 520  
;; TYPE: DNA  
;; ORGANISM: Homo Sapiens  
US-11-266-748A-44240

Query Match 20.8%; Score 514; DB 8; Length 520;  
Best Local Similarity 100.0%; Pred. No. 4.3e-125; Indels 0; Gaps 0;  
Matches 514; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1956 CTGAAGTGTCAACACAGTCAGATTCTGGGAAGTGACAAATGCAATTTGCAATTTAGAAGTG 2015  
DB 1 CTGAAGTGTCAACACAGTCAGATTCTGGGAAGTGACAAATGCAATTTGCAATTTAGAAGTG 60  
QY 2016 TGCCAGCTTCAAAAACCCGATGTGGGTATCAGGAGCAGCGGCTTAAGGCGATGCC 2075  
DB 61 TGCCAGCTTCAAAAACCCGATGTGGGTATCAGGAGCAGCGGCTTAAGGCGATGCC 120  
QY 2076 TGGGTTTACAAGATTAGTGGAGACATCCTATCTAGCTGCAAGGTATAATTTGATGGAT 2135  
DB 121 TGGGTTTACAAGATTAGTGGAGACATCCTATCTAGCTGCAAGGTATAATTTGATGGAT 180  
QY 2136 TCTTCATCTCTGCGGATGAGTGTGGGTGTGATACAGCCTACATAAAGACTGTTATGATC 2195  
DB 181 TCTTCATCTCTGCGGATGAGTGTGGGTGTGATACAGCCTACATAAAGACTGTTATGATC 240  
QY 2196 GCTTTGATTTAAAGTTTCATGGAACTCAACCTGTTTCTAAAGAGCTATCTTAAGACC 2255  
DB 241 GCTTTGATTTAAAGTTTCATGGAACTCAACCTGTTTCTAAAGAGCTATCTTAAGACC 300  
QY 2256 AATATCTCTTTGTTTAAACAAAGATATTTTGTGTATGAATCTAAATCAAGCCCA 2315  
DB 301 AATATCTCTTTGTTTAAACAAAGATATTTTGTGTATGAATCTAAATCAAGCCCA 360  
QY 2316 TCTGTCAATATGTACTGTCTTTTAAATCAATGTGTTTGTATATTAATAATTTGTGAC 2375  
DB 361 TCTGTCAATATGTACTGTCTTTTAAATCAATGTGTTTGTATATTAATAATTTGTGAC 420  
QY 2376 TTTCTTAGATTCATTTCCATATGTGAATGAAGCTCTTAAGTATGCTCTTTGTAATGTG 2435  
DB 421 TTTCTTAGATTCATTTCCATATGTGAATGAAGCTCTTAAGTATGCTCTTTGTAATGTG 480  
QY 2436 TAATTTCTTTTGAATAAACCATTGTTGAATA 2469  
DB 481 TAATTTCTTTTGAATAAACCATTGTTGAATA 514

RESULT 15  
US-11-145-471-52  
;; Sequence 52, Application US/11145471  
;; Publication No. US20060147947A1  
;; GENERAL INFORMATION:  
;; APPLICANT: Apfeld, Javier  
;; APPLICANT: O'Conner, Gregory  
;; TITLE OF INVENTION: AMPK Pathway Components  
;; FILE REFERENCE: 13407-0570011

;; CURRENT APPLICATION NUMBER: US/11/145,471  
;; CURRENT FILING DATE: 2005-06-03  
;; PRIOR APPLICATION NUMBER: PCT/US/2003/38628  
;; PRIOR FILING DATE: 2003-12-04  
;; PRIOR APPLICATION NUMBER: US 60/430,804  
;; PRIOR FILING DATE: 2002-12-04  
;; PRIOR APPLICATION NUMBER: US 60/488,261  
;; PRIOR FILING DATE: 2003-07-18  
;; PRIOR APPLICATION NUMBER: 60/578,804  
;; PRIOR FILING DATE: 2004-06-10  
;; NUMBER OF SEQ ID NOS: 56  
;; SOFTWARE: PatentIn version 3.2  
;; SEQ ID NO 52  
;; LENGTH: 1863  
;; TYPE: DNA  
;; ORGANISM: Homo sapiens  
US-11-145-471-52

Query Match 8.4%; Score 208.4; DB 9; Length 1863;  
Best Local Similarity 55.6%; Pred. No. 2.3e-44;  
Matches 470; Conservative 0; Mismatches 361; Indels 15; Gaps 3;  
QY 212 TGAAACTATTGGGACAGGTGGCTTTGCAAGGTCAAACTTGGCTGCCATATCTTACTCG 271  
DB 86 TGACAGCTGGGGTGGCACCCTTCGGCAAGTGAAGTTGGCAACATGAATTTGACTGG 145  
QY 272 AGAGATGGTAGCTATAAAAAATCATGGATAAAACACACACTAGGGAG-----TGATTTGCC 325  
DB 146 GCATAAAGTAGCTGTGAAGATACTCAATCGACAGAAGATTTCGGAGCCTTGATGTGGTAGS 205  
QY 326 CCGGATCAAAACGGAGATTGAGGCCCTTGAGAACCTTGAGACATCAGACATATATGTCAACT 385  
DB 206 AAAAAATCCGACAGAAATTCAGAACCTCAAGCTTTTCAGGSCATCTCATATAATTAACCT 265  
QY 386 CTACCATGTGTAGAGACAGCCCAACAAATATTCATGGTTCTTGAGTACTGCCCTCGAGG 445  
DB 266 GTACCAAGTCACTCAGTACACCATCTGATATTTTCATGGTGTGGAATATGTCTCAGGAG 325  
QY 446 AGAGCTGTTTGGTACTATATAATTTCCAGGATCGCCTGTGAGAGAGGAGACCCGGGTTGT 505  
DB 326 AGAGCTATTTGATTATATCTGTAAGAAATGGAAGGCTGGATGAAAGAAAGTCGGGTCT 385  
QY 506 CTTCCGTGATAGTATCTGCTGTTGCTTATGTGACAGCCAGGGCTATGCTCAGAGGA 565  
DB 386 GTTCCAACAGATCTCTTCTGGTGTGATTAATTTGTCACAGGCATATGTTGTTCCATAGAGA 445  
QY 566 CTTCAAGCCAGAAATTTGCTGTTTGTGATGAATCATATAATTTAAAGCTGATTGACTTTCG 625  
DB 446 TTTGAAACCTGAAATGCTCTGTTGATGCAACATGAATGCAAAAGATAGCTGATTTTGG 505  
QY 626 TCTCTGTGCAAAACCCAGGGTAAACAAGGATTTACCATCTACAGACATGCTGGGAGTCT 685  
DB 506 TCTTTC-----AAACATGATGTGAGTGGTGAATTTTAAAGACAAAGTTGGCTCACC 559  
QY 686 GGCCTATGAGCACCTGAGTTTAATCAAGGCAAAATCATATCTTTGGATCAGAGGAGATGT 745  
DB 560 CAACTATGTGCACCAAGTAATTTTCAGGAAGATTGTATGCAAGGCCACAGAGGTAGATAT 619  
QY 746 TTGGAGCATGGCATACTGTTATATGTTCTTATGTGTGATTTCTACCAATTTGATGATGA 805  
DB 620 ATGGAGCATGGGGTTATTTCTCTATGCTTTATATATGTGGAACCTTTCCATTTGATGATGA 679  
QY 806 TAATGTAATGGCTTTATACAAGAATTTATGAGAGGAAATATGATGTTTCCCAAGTGGCT 865  
DB 680 CCATGTGCCAATCTTTTAAAGAGATATGTGATGGGATCTTCTATACCCCTCAATATTT 739  
QY 866 CTCCTCCAGTAGCATCTGCTTTCTTCAACAAATGTGCAAGGTGGACCCCAAGAAACGGAT 925  
DB 740 AAATCCTTCTGTGATTAGCCTTTTGAACATATGCTGCAAGTGGATCCCATGAAGGGGC 799  
QY 926 TTCTATGAAATCTTATGAAACCATCCCTGGATCATGCAAGA---TTACAACTATCTCTGT 982  
DB 800 CTCAATCAAGATATCAGGGAACATGAATGGTTTAAACAGGACCTTCCAAAAATATCTCTT 859

Qy 983 TGAGTGGCAAGCAGAAATCCTTTTATTCACCTCGATGATGCGTAACAGAACTTTC 1042  
Db 860 TCCTGAGGATCCATCATATAGTTCAACCATGATGATGAAGCCTTAAAGAGATG 919

Qy 1043 TGTACA 1048  
Db 920 TGAATA 925

Search completed: November 21, 2006, 23:04:41  
Job time : 561 secs

## Protein Sequence Searches - February 2005

All of the sequence databases on ABSS have recently been updated.

- Please note that the curators of the UniProt database have purged some temporary accession numbers from the most recent version of UniProt. These sequences have been assigned new permanent accession numbers. The new UniProt record may not contain the previous temporary accession number.
- If you encounter an accession number from an older search run against UniProt (results file extension **.rup**) that can no longer be found in the database, the permanent record with the new accession number can be found by searching the old accession number in the UniProt Protein Archive database (UniPARC) at:

<http://www.pir.uniprot.org/database/archive.shtml>

If you have any questions regarding this information or your results, please contact any STIC searcher.

**When submitting sequence search results for scanning into IFW, please include a copy of this attachment to assist any future Examiners or members of the public who may encounter UniProt temporary accession numbers.**

**This Page Blank (uspto)**



November 2005

Published\_Applications Nucleic Acid and Published\_Applications Amino Acid database searches now generate two sets of results each. The Published\_Applications databases have been split into two parts to reduce the amount of time required for their daily updates. This results in more machine time being available for processing searches.

Newly published applications will appear in the Published\_Applications New databases; older published applications make up the Published\_Applications Main databases.

Searches run against Nucleic Acid Published\_Applications produce two sets of results, with the extensions **.rnpbm** (Published\_Applications\_NA\_Main) and **.rnpbn** (Published\_Applications\_NA\_New).

Searches run against Amino Acid Published\_Applications produce two sets of results, with the extensions **.rapbm** (Published\_Applications\_AA\_Main) and **.rapbn** (Published\_Applications\_AA\_New).

**this Page Blank (uspto)**

GenCore version 5.1.9  
 Copyright (c) 1993 - 2006 Bioceleration Ltd.  
 OM protein - protein search, using sw model  
 Run on: November 21, 2006, 14:04:56 ; Search time 308 Seconds  
 (without alignments)  
 1955.147 Million cell updates/sec  
 Title: US-10-656-598-2  
 Perfect score: 3447  
 Sequence: 1 MKDYDELLKYYELHETIGT.....KGDANVYKRLVEDILSSCKV 651

Scoring table: BLOSUM62  
 Gapop 10.0 , Gapext 0.5  
 Searched: 2849598 seqs, 925015592 residues  
 Total number of hits satisfying chosen parameters: 2849598  
 Minimum DB seq length: 0  
 Maximum DB seq length: 2000000000  
 Post-processing: Minimum Match 0%  
 Maximum Match 100%  
 Listing first 45 summaries

Database : Uniprot 7.2.2\*  
 1: uniprot\_sprot:\*  
 2: uniprot\_trembl:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	3447	100.0	651	1	MELK_HUMAN	Q14680 homo sapien
2	3443	99.9	650	2	Q5T263_HUMAN	Q5t263 homo sapien
3	3443	99.5	651	2	Q53GX0_HUMAN	Q53gx0 homo sapien
4	2878	83.5	643	2	Q3UMB0_MOUSE	Q3umb0 mus musculus
5	2864	83.1	643	1	MELK_MOUSE	Q61846 mus musculus
6	2863	83.1	643	2	Q3TUF3_MOUSE	Q3tuf3 mus musculus
7	2863	83.1	643	2	Q3TPU1_MOUSE	Q3tpu1 m 12 days e
8	2859	82.9	643	2	Q5U440_MOUSE	Q5u440 mus musculus
9	2852	82.7	643	2	Q3TU15_MOUSE	Q3tu15 mus musculus
10	2348	68.1	457	2	Q4R6B9_MACFA	Q4r6b9 macaca fasc
11	2228	64.6	451	2	Q91821_XENLA	Q91821 xenopus lae
12	2083	60.4	657	2	Q5ZL85_CHICK	Q5zl85 gallus gall
13	1920	55.7	676	2	Q7Z2N5_BRARE	Q7z2n5 brachydanio
14	1911	55.4	676	2	Q7ZU72_BRARE	Q7zu72 brachydanio
15	1528.5	44.3	476	2	Q498M1_XENLA	Q498m1 xenopus lae
16	1307.5	37.9	689	2	Q4RK91_TETNG	Q4rk91 tetraodon n
17	1187.5	34.5	688	2	Q95UF4_ANCCA	Q95uf4 ancylostoma
18	1156	33.5	726	2	Q9UAY1_CAEEL	Q9uay1 caenorhabdi
19	1128.5	32.7	701	2	Q624C7_CAEHR	Q624c7 caenorhabdi
20	1071.5	31.1	722	2	Q86FM0_CAEEL	Q86fm0 caenorhabdi
21	804.5	23.3	729	2	Q5U5B2_XENLA	Q5u5b2 xenopus lae
22	792.5	23.0	753	1	MARX3_MOUSE	Q03141 mus musculus
23	791	22.9	744	2	Q9JKE5_MOUSE	Q9jke5 mus musculus
24	789.5	22.9	783	2	Q5BL77_XENTR	Q5bl77 xenopus tro
25	788.5	22.9	797	1	MARX3_RAT	Q8vho0 rattus norv
26	780.5	22.6	780	2	Q804T1_XENLA	Q804t1 xenopus lae
27	778.5	22.6	725	2	Q804T2_XENLA	Q804t2 xenopus lae
28	775.5	22.5	776	2	Q7ZYL7_XENLA	Q7zyl7 xenopus lae
29	775	22.5	785	2	Q8QGV3_XENLA	Q8qgv3 xenopus lae
30	767.5	22.3	788	1	MARX2_HUMAN	Q7kx17 homo sapien
31	766.5	22.2	722	2	Q802W0_BRARE	Q802w0 brachydanio

RESULT 1  
 ID MELK\_HUMAN STANDARD; PRT; 651 AA.  
 AC Q14680; Q7L3C3;  
 DT 19-JUL-2004, integrated into UniProtKB/Swiss-Prot.  
 DT 19-JUL-2004, sequence version 3.  
 DE Maternal embryonic leucine zipper kinase (EC 2.7.1.37) (hMELK)  
 DE (Protein kinase PK38) (hPK38).  
 GN Name=MELK; Synonyms=KIAA0175;  
 OS Homo sapiens (Human)  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;  
 OC Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA], AND TISSUE SPECIFICITY.  
 RC TISSUE=Bone marrow;  
 RX MEDLINE=96281124; PubMed=8724849; DOI=10.1093/dnares/3.1.17;  
 RA Nagase T., Seki N., Ishikawa K., Tanaka A., Nomura N.;  
 RT "Prediction of the coding sequences of unidentified human genes. V.  
 RT The coding sequences of 40 new genes (KIAA0161-KIAA0200) deduced by  
 RT analysis of cDNA clones from human cell line KG-1.";  
 RL DNA Res. 3:17-24 (1996).  
 RN [2]  
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].  
 RC TISSUE=Colon;  
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Schuler G.D.,  
 RA Klausner R.D., Collins P.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Udgin T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Raha S.A., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,  
 RA Bosak S.A., McSwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalón D.K., Muzny D.M., Sodergren B.J., Lu X., Gibbs R.A.,  
 RA Fahey J., Helton E., Kettelman M., Maman A., Rodriguez S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
 RA Butterfield Y.S.N., Krzywinski M.I., Skalek U., Smallos D.E.,  
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length human  
 RT and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).  
 RN [3]  
 RP INTERACTION WITH ZNF622, AND PHOSPHORYLATION OF ZNF622.  
 RC TISSUE=Keratinocyte;  
 RX PubMed=11802789; DOI=10.1042/0264-6021.3610597;  
 RA Seong H.-A., Gil M., Kim K.-T., Kim S.-J., Ha H.;  
 RT "Phosphorylation of a novel zinc-finger-like protein, ZPR9, by murine

## ALIGNMENTS

RT protein serine/threonine kinase 38 (MPX38).";  
RN Biochem. J. 361:597-604(2002).  
[4]  
RP INTERACTION WITH PPP1R8, AUTOPHOSPHORYLATION, MUTAGENESIS OF ASP-150;  
RP THR-345; THR-387; THR-409; THR-415; THR-428; THR-446; THR-460;  
RP THR-466; THR-478 AND THR-518, PHOSPHORYLATION SITE THR-478, AND  
RP FUNCTION.  
RX PubMed=14699119; DOI=10.1074/jbc.M111466200;  
RA Vulsteke V., Beullens M., Boudrez A., Keppens S., Van Eynde A.,  
RA Rider M.H., Stalmans W., Bollen M.;  
RT "Inhibition of spliceosome assembly by the cell cycle-regulated  
RT protein kinase MELK and involvement of splicing factor NIPP1.";  
RL J. Biol. Chem. 279:8642-8647(2004).  
CC -i- FUNCTION: Phosphorylates ZNF622 and may contribute to its  
CC redirection to the nucleus. May be involved in the inhibition of  
CC spliceosome assembly during mitosis.  
CC -i- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.  
CC -i- SUBUNIT: Interacts with ZNF622 and PPP1R8.  
CC -i- SUBCELLULAR LOCATION: Cytoplasm (Potential).  
CC -i- TISSUE SPECIFICITY: Expressed in placenta, kidney, thymus, testis,  
CC ovary and intestine.  
CC -i- PTM: Autophosphorylated. Thr-478 phosphorylation during mitosis  
CC promotes interaction with PPP1R8 (Probable).  
CC -i- SIMILARITY: Belongs to the Ser/Thr protein kinase family. SNF1  
CC subfamily.  
CC -i- SIMILARITY: Contains 1 KAI (kinase-associated) domain.  
CC  
CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>  
CC Distributed under the Creative Commons Attribution-NoDerivs License  
CC  
CC -----  
CC EMBL; BC014039; AAH14039.1; -; mRNA.  
CC HSSP; P31751; 1GZK.  
CC Ensembl; ENSG00000165304; Homo sapiens.  
CC HGNC; HGNC:16870; MELK.  
CC MIM; 607025; gene.  
CC InterPro; IPR001772; Kinase C.  
CC InterPro; IPR000719; Prot kinase.  
CC InterPro; IPR008271; Ser\_Ehr\_pkin\_AS.  
CC InterPro; IPR002290; Ser\_thr\_pkinase.  
CC InterPro; IPR001245; Tyr\_pkinase.  
CC Pfam; PF02149; KAI; 1.  
CC Pfam; PF00069; Pkinase; 1.  
CC ProDom; PD000001; Prot kinase; 1.  
CC SMART; SM00220; S\_TKC\_1.  
CC PROSITE; PS00107; PROTEIN KINASE ATP; 1.  
CC PROSITE; PS00011; PROTEIN KINASE DOM; 1.  
CC PROSITE; PS00108; PROTEIN KINASE ST; 1.  
CC ATP-binding; Kinase; Nucleotide-binding; Phosphorylation;  
KW Serine/threonine-protein kinase; Transferase.  
FT CHAIN 1 651  
FT /FTID=PRO\_0000086323.  
FT Protein kinase.  
FT DOMAIN 11 263  
FT DOMAIN 602 651  
FT KAI.  
FT NP\_BIND 17 25  
FT ACT\_SITE 132 132  
FT BINDING 40 40  
FT MOD\_RES 478 478  
FT MUTAGEN 150 150  
FT T-A: No effect on interaction with  
FT PPP1R8.  
FT T-A: No effect on interaction with  
FT PPP1R8.  
FT T-A: No effect on interaction with  
FT PPP1R8.  
FT T-A: No effect on interaction with  
FT PPP1R8.  
FT T-A: No effect on interaction with  
FT PPP1R8.  
FT T-A: Inhibits interaction with PPP1R8.  
FT T-A: Inhibits interaction with PPP1R8.  
FT T-A: Inhibits interaction with PPP1R8.  
FT T-A: Strongly inhibits interaction with

FT MUTAGEN 518 518  
FT T-A: Enhances enzymatic activity.  
FT T-A: No effect on interaction with  
FT PPP1R8.  
SQ SEQUENCE 651 AA; 74642 MW; 57F05CDC6122B570 CRC64;  
Query Match 100.0%; Score 3447; DB 1; Length 651;  
Best Local Similarity 100.0%; Pred. No. 6.2e-224;  
Matches 651; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MKDYDELLKYVYELHETIGTGGFAKVKLACHILITGEMVAIKIMDKNTGLSGDLPRIKTEIRA 60  
DB 1 MKDYDELLKYVYELHETIGTGGFAKVKLACHILITGEMVAIKIMDKNTGLSGDLPRIKTEIRA 60  
QY 61 LKNLRHQHICQLYHYVLETANKIFMVLEYCPGGELEFDYIISQDRLSEETRVVFRQIVSAV 120  
DB 61 LKNLRHQHICQLYHYVLETANKIFMVLEYCPGGELEFDYIISQDRLSEETRVVFRQIVSAV 120  
QY 121 AYVHSGYVADYVSMGILLYVLMCGFLPDDNNVWALYKKIMRGKYDVDPKWLSPSSILL 240  
DB 121 AYVHSGYVADYVSMGILLYVLMCGFLPDDNNVWALYKKIMRGKYDVDPKWLSPSSILL 240  
QY 181 QGKSYLGSEADYVSMGILLYVLMCGFLPDDNNVWALYKKIMRGKYDVDPKWLSPSSILL 300  
DB 181 QGKSYLGSEADYVSMGILLYVLMCGFLPDDNNVWALYKKIMRGKYDVDPKWLSPSSILL 300  
QY 241 QQMLQVDPKKRISMKNLLNHPIMQDYNYPVWQSKNPFILHDDDCVTELSVHHRNRQT 300  
DB 241 QQMLQVDPKKRISMKNLLNHPIMQDYNYPVWQSKNPFILHDDDCVTELSVHHRNRQT 300  
QY 301 MEDLISLWQYDHLTATYLLLLAKKARGKPVRLRLSFSGQASATPFTDIKNNWSLEDV 360  
DB 301 MEDLISLWQYDHLTATYLLLLAKKARGKPVRLRLSFSGQASATPFTDIKNNWSLEDV 360  
QY 361 TASDKNYVAGLIDYDWCEDDLSTGAATPRTSFTKYWTESNGVESKSLTPALCRTPANKL 420  
DB 361 TASDKNYVAGLIDYDWCEDDLSTGAATPRTSFTKYWTESNGVESKSLTPALCRTPANKL 420  
QY 421 KKNENYVTPKSAVKNEEYFMFPPEKTPVKNQKHREILTPNRYTTPSKARNOCLKETPI 480  
DB 421 KKNENYVTPKSAVKNEEYFMFPPEKTPVKNQKHREILTPNRYTTPSKARNOCLKETPI 480  
QY 481 KIPVNSTGTDKLMTGVISPERCSRVELDLNQAHEETPKRGAKVFGSLERGLDKVITV 540  
DB 481 KIPVNSTGTDKLMTGVISPERCSRVELDLNQAHEETPKRGAKVFGSLERGLDKVITV 540  
QY 541 LTRSKKSGSARDGPRRLKLUHNTVITRLVNPOLLNEIMSLIPKKHVDVFOKGYTLKCO 600  
DB 541 LTRSKKSGSARDGPRRLKLUHNTVITRLVNPOLLNEIMSLIPKKHVDVFOKGYTLKCO 600  
QY 601 QSDFGKVTMQFLEVCQLQKPDVVGIRRQRLKGDWVYKRLVEDILSSCKV 651  
DB 601 QSDFGKVTMQFLEVCQLQKPDVVGIRRQRLKGDWVYKRLVEDILSSCKV 651  
RESULT 2  
Q5T263 HUMAN  
ID Q5T263.HUMAN PRELIMINARY; PRT; 650 AA.  
AC Q5T263.  
DT 21-DEC-2004, integrated into UniProtKB/TrEMBL.  
DT 21-DEC-2004, sequence version 1.  
DT 07-FEB-2006, entry version 10.  
DE Maternal embryonic leucine zipper kinase (Fragment).  
GN Name=MELK; ORFNames=RP11-8N6.1-001;  
OS Homo sapiens (Human)  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;  
OC Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP NUCLEOTIDE SEQUENCE.  
RA Beasley H.;  
RL Submitted (MAY-2005) to the EMBL/GenBank/DBJ databases.  
RN [2]

RP NUCLEOTIDE SEQUENCE.  
RA Garner P.;  
RL Submitted (MAY-2005) to the EMBL/GenBank/DBJ databases.  
CC -1- SIMILARITY: Belongs to the Ser/Thr protein kinase family.  
CC -----  
CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>  
CC Distributed under the Creative Commons Attribution-NoDerivs License  
CC -----  
DR EMBL; AL442063; CAIL1034.1; JOINED; Genomic DNA.  
DR EMBL; AL442063; CAIL16995.1; -; Genomic DNA.  
DR EMBL; AL354932; CAIL16995.1; JOINED; Genomic DNA.  
DR EMBL; AL354932; CAIL1034.1; -; Genomic DNA.  
DR Ensembl; ENSG00000165304; Homo sapiens.  
DR GO; GO:0005524; P:ATP binding; IEA.  
DR GO; GO:000166; F:nucleotide binding; IEA.  
DR GO; GO:0004674; P:protein serine/threonine kinase activity; IEA.  
DR GO; GO:0016740; F:transferase activity; IEA.  
DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.  
DR InterPro; IPR001772; Kinase.C.  
DR InterPro; IPR000719; Prot\_kinase.  
DR InterPro; IPR008271; Ser\_thr\_pkin\_AS.  
DR InterPro; IPR002290; Ser\_thr\_pkinase.  
DR InterPro; IPR001245; Tyr\_pkinase.  
DR Pfam; PF02149; KAI; 1.  
DR Pfam; PF00069; Pkinase; 1.  
DR ProDom; PD000001; Prot\_kinase; 1.  
DR SMART; SM00220; S\_TKC; 1.  
DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; UNKNOWN\_1.  
DR PROSITE; PS00011; PROTEIN\_KINASE\_DOM; 1.  
DR PROSITE; PS00108; PROTEIN\_KINASE\_ST; UNKNOWN\_1.  
KW ATP-binding; Kinase; Nucleotide-binding;  
KW Serine/threonine-protein kinase; Transferase.  
FT NON\_TER 650  
SQ SEQUENCE 650 AA; 74543 MW; 705CDC6122E5703E CRC64;

Query Match 99.9%; Score 3443; DB 2; Length 650;  
Best Local Similarity 100.0%; Pred. No. 1.2e-223;  
Matches 650; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKDYDELLKYYELHETIGTGPAKVKLACHILGEMVAIKIMDKNTLGSPLPRIKTEIEA 60  
DB 1 MKDYDELLKYYELHETIGTGPAKVKLACHILGEMVAIKIMDKNTLGSPLPRIKTEIEA 60

QY 61 LKVLRHQICQLVHVLETANKIPWLEYCPGBELFYIISQDRLSEETRVVFRQIVSAV 120  
DB 61 LKVLRHQICQLVHVLETANKIPWLEYCPGBELFYIISQDRLSEETRVVFRQIVSAV 120

QY 121 AVHSGYAHRLDKPENLLPDEYHKLKIDFGICAKPKGNKDYHLOTCCGSLAYAAPELI 180  
DB 121 AVHSGYAHRLDKPENLLPDEYHKLKIDFGICAKPKGNKDYHLOTCCGSLAYAAPELI 180

QY 181 QGKSYLGSEADVMSGILLYVLMCGFLPFDDNNMALYKKIMRGKYDVPKWLSPSSILL 240  
DB 181 QGKSYLGSEADVMSGILLYVLMCGFLPFDDNNMALYKKIMRGKYDVPKWLSPSSILL 240

QY 241 QQMLQVDPKKRISMKNLNHPWTIMQDYNYPVEWQSKNPFFHLDDCCVTELSVHHRNRQT 300  
DB 241 QQMLQVDPKKRISMKNLNHPWTIMQDYNYPVEWQSKNPFFHLDDCCVTELSVHHRNRQT 300

QY 301 MEDLISLWQYDHLTATYLLILLAKARGKPVRLSLSSFCGQASATPTDIKSNWLSLEDV 360  
DB 301 MEDLISLWQYDHLTATYLLILLAKARGKPVRLSLSSFCGQASATPTDIKSNWLSLEDV 360

QY 361 TASDKNVAAGLIIDYWCEDDLSTGAATPRTSQFTKYWTESNGVESKSLTPALCRTPANKL 420  
DB 361 TASDKNVAAGLIIDYWCEDDLSTGAATPRTSQFTKYWTESNGVESKSLTPALCRTPANKL 420

QY 421 KKNENYTPKSAVKNEEYFMFPPEKTPVKNQKHREILITTPNRYTTPSKARNQCLKETPI 480  
DB 421 KKNENYTPKSAVKNEEYFMFPPEKTPVKNQKHREILITTPNRYTTPSKARNQCLKETPI 480

QY 481 KIPVNSTGTDKLTMTGTVISPERRCRSEVELDLNOAHMEETPKRGAKVFGSLERGLDKVITV 540

Db 481 KIPVNSTGTDKLTMTGTVISPERRCRSEVELDLNOAHMEETPKRGAKVFGSLERGLDKVITV 540

QY 541 LTRSKRGSGARDGPRRLKLYHNVTTTTLVNPQLLNEIMSLIPKQHVDFVQGYTLKCO 600  
Db 541 LTRSKRGSGARDGPRRLKLYHNVTTTTLVNPQLLNEIMSLIPKQHVDFVQGYTLKCO 600

QY 601 QSDFGKVTWQFLEVCQLOKQPDVVGIRRRORLKGDAWVYKRLVEDILSSCK 650  
Db 601 QSDFGKVTWQFLEVCQLOKQPDVVGIRRRORLKGDAWVYKRLVEDILSSCK 650

RESULT 3  
Q53GX0\_HUMAN PRELIMINARY; PRT; 651 AA.  
AC Q53GX0;  
DT 24-MAY-2005, integrated into UniProtKB/TrEMBL.  
DT 24-MAY-2005, sequence version 1.  
DE 07-PEB-2006, entry version 6.  
DE Maternal embryonic leucine zipper kinase variant (Fragment).  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;  
OC Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP NUCLEOTIDE SEQUENCE.  
RC TISSUE=Liver;  
RA MEDLINE=94171032; PubMed=8125298; DOI=10.1016/0378-1119(94)90802-8;  
RA Maruyama K., Sugano S.;  
RT "Oligo-capping: a simple method to replace the cap structure of  
RT eucaryotic mRNAs with oligoribonucleotides.";  
RL Gene 138:171-174(1994).  
RN [2]  
RP NUCLEOTIDE SEQUENCE.  
RC TISSUE=Liver;  
RX MEDLINE=98038986; PubMed=9373149; DOI=10.1016/S0378-1119(97)00411-3;  
RA Suzuki Y., Yoshitomo K., Maruyama K., Sugama A., Sugano S.;  
RL Submitted (APR-2005) to the EMBL/GenBank/DBJ databases.  
CC -1- SIMILARITY: Belongs to the Ser/Thr protein kinase family.  
CC -----  
CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>  
CC Distributed under the Creative Commons Attribution-NoDerivs License  
CC -----  
DR EMBL; AK222811; BAD96531.1; -; mRNA.  
DR Ensembl; ENSG00000165304; Homo sapiens.  
DR GO; GO:0005524; P:ATP binding; IEA.  
DR GO; GO:000166; F:nucleotide binding; IEA.  
DR GO; GO:0004674; P:protein serine/threonine kinase activity; IEA.  
DR GO; GO:0016740; F:transferase activity; IEA.  
DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.  
DR InterPro; IPR001772; Kinase.C.  
DR InterPro; IPR000719; Prot\_kinase.  
DR InterPro; IPR008271; Ser\_thr\_pkin\_AS.  
DR InterPro; IPR002290; Ser\_thr\_pkinase.  
DR InterPro; IPR001245; Tyr\_pkinase.  
DR Pfam; PF02149; KAI; 1.  
DR Pfam; PF00069; Pkinase; 1.  
DR ProDom; PD000001; Prot\_kinase; 1.  
DR SMART; SM00220; S\_TKC; 1.  
DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; 1.  
DR PROSITE; PS00011; PROTEIN\_KINASE\_DOM; 1.  
DR PROSITE; PS00108; PROTEIN\_KINASE\_ST; UNKNOWN\_1.  
KW ATP-binding; Kinase; Nucleotide-binding;  
KW Serine/threonine-protein kinase; Transferase.  
FT NON\_TER 1



RA Verardo R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C.,  
RA Wilmig Z.G., Wynshaw-Boris A., Yanagisawa M., Yang I., Yang L.,  
RA Yuan Z., Zavanon M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,  
RA Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,  
RA Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,  
RA Hara A., Hashizume M., Inotani K., Ishii Y., Itoh M., Kagawa I.,  
RA Miyazaki A., Sakai K., Saeki D., Shibata K., Shinagawa A.,  
RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,  
RA Birney E., Hayashizaki Y.,  
RT "Analysis of the mouse transcriptome based on functional annotation of  
RT 60,770 full-length cDNAs";  
RL Nature 420:563-573 (2002).  
RN [5]  
RP NUCLEOTIDE SEQUENCE.  
RC TISSUE=Mammary gland;  
RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;  
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,  
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,  
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,  
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,  
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,  
RA Fleischnann W., Gaasterland T., Gissi C., King B., Kochiwa H.,  
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,  
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,  
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,  
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.P.,  
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,  
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,  
RA Lyons P., Marchionni L., Maehima J., Mazzarelli J., Mombaerts P.,  
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,  
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,  
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,  
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,  
RA Hayashizaki Y.,  
RT "Functional annotation of a full-length mouse cDNA collection";  
RL Nature 409:685-690 (2001).  
RN [6]  
RP NUCLEOTIDE SEQUENCE.  
RC TISSUE=Mammary gland;  
RX MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;  
RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,  
RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.,  
RT "Normalization and subtraction of cap-trapper-selected cDNAs to  
RT prepare full-length cDNA libraries for rapid discovery of new genes";  
RL Genome Res. 10:1617-1630 (2000).  
RN [7]  
RP NUCLEOTIDE SEQUENCE.  
RC TISSUE=Mammary gland;  
RX MEDLINE=20530913; PubMed=11076961; DOI=10.1101/gr.152600;  
RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,  
RA Konno H., Akiyama J., Nishi K., Kitsuunai T., Tashiro H., Itoh M.,  
RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,  
RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,  
RA Fujiwaka S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,  
RA Itoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuura S., Kawai J.,  
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.,  
RT "RIKEN integrated sequence analysis (RISA) system-384-format  
RT sequencing pipeline with 384 multipillar sequencer";  
RL Genome Res. 10:1757-1771 (2000).  
RN [8]  
RP NUCLEOTIDE SEQUENCE.  
RC TISSUE=Mammary gland;  
RX MEDLINE=20530913; PubMed=11076961; DOI=10.1101/gr.152600;  
RA Arakawa T., Carninci P., Fukuda S., Hashizume W., Hayashida K.,  
RA Hori F., Iida J., Imamura K., Inotani K., Itoh M., Kanagawa S.,  
RA Kawai J., Kojima M., Konno H., Murata M., Nakamura M., Ninomiya N.,  
RA Nishiyori H., Nomura K., Ohno M., Sakazume N., Sano H., Sasaki D.,  
RA Shibata K., Shiraki T., Tagami M., Tagami Y., Waki K., Watahiki A.,  
RA Muramatsu M., Hayashizaki Y.,  
RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.  
CC -1- SIMILARITY: Belongs to the Ser/Thr protein kinase family.  
CC  
CC Copyrighted by the Uniprot Consortium, see <http://www.uniprot.org/terms>  
CC Distributed under the Creative Commons Attribution-NoDerivs License

CC EMBL; AK145021; BAE26188.1; -, mRNA.  
DR MGI; MGI:106924; Melk.  
DR GO; GO:0005713; F:cytoplasm; IDA.  
DR GO; GO:0004713; F:proteolysis; kinase activity; RCA.  
DR Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,  
DR Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,  
DR Hara A., Hashizume M., Inotani K., Ishii Y., Itoh M., Kagawa I.,  
DR Miyazaki A., Sakai K., Saeki D., Shibata K., Shinagawa A.,  
DR Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,  
DR Birney E., Hayashizaki Y.,  
RT "Analysis of the mouse transcriptome based on functional annotation of  
RT 60,770 full-length cDNAs";  
RL Nature 420:563-573 (2002).  
RN [5]  
RP NUCLEOTIDE SEQUENCE.  
RC TISSUE=Mammary gland;  
RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;  
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,  
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,  
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,  
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,  
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,  
RA Fleischnann W., Gaasterland T., Gissi C., King B., Kochiwa H.,  
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,  
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,  
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,  
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.P.,  
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,  
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,  
RA Lyons P., Marchionni L., Maehima J., Mazzarelli J., Mombaerts P.,  
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,  
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,  
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,  
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,  
RA Hayashizaki Y.,  
RT "Functional annotation of a full-length mouse cDNA collection";  
RL Nature 409:685-690 (2001).  
RN [6]  
RP NUCLEOTIDE SEQUENCE.  
RC TISSUE=Mammary gland;  
RX MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;  
RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,  
RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.,  
RT "Normalization and subtraction of cap-trapper-selected cDNAs to  
RT prepare full-length cDNA libraries for rapid discovery of new genes";  
RL Genome Res. 10:1617-1630 (2000).  
RN [7]  
RP NUCLEOTIDE SEQUENCE.  
RC TISSUE=Mammary gland;  
RX MEDLINE=20530913; PubMed=11076961; DOI=10.1101/gr.152600;  
RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,  
RA Konno H., Akiyama J., Nishi K., Kitsuunai T., Tashiro H., Itoh M.,  
RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,  
RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,  
RA Fujiwaka S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,  
RA Itoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuura S., Kawai J.,  
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.,  
RT "RIKEN integrated sequence analysis (RISA) system-384-format  
RT sequencing pipeline with 384 multipillar sequencer";  
RL Genome Res. 10:1757-1771 (2000).  
RN [8]  
RP NUCLEOTIDE SEQUENCE.  
RC TISSUE=Mammary gland;  
RX MEDLINE=20530913; PubMed=11076961; DOI=10.1101/gr.152600;  
RA Arakawa T., Carninci P., Fukuda S., Hashizume W., Hayashida K.,  
RA Hori F., Iida J., Imamura K., Inotani K., Itoh M., Kanagawa S.,  
RA Kawai J., Kojima M., Konno H., Murata M., Nakamura M., Ninomiya N.,  
RA Nishiyori H., Nomura K., Ohno M., Sakazume N., Sano H., Sasaki D.,  
RA Shibata K., Shiraki T., Tagami M., Tagami Y., Waki K., Watahiki A.,  
RA Muramatsu M., Hayashizaki Y.,  
RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.  
CC -1- SIMILARITY: Belongs to the Ser/Thr protein kinase family.  
CC  
CC Copyrighted by the Uniprot Consortium, see <http://www.uniprot.org/terms>  
CC Distributed under the Creative Commons Attribution-NoDerivs License

RESULT 5

MELK\_MOUSE

ID MELK\_MOUSE

STANDARD;

PRT; 643 AA.



AC Q61846; Q61804; Q6ZQH6;  
 DT 19-JUL-2004, integrated into UniProtKB/Swiss-Prot.  
 DT 01-NOV-1996, sequence version 1.  
 DT 07-MAR-2006, entry version 42.  
 DE Maternal embryonic leucine zipper kinase (EC 2.7.1.37) (Protein kinase  
 DE PK38) (mpk38).  
 GN Name:Melk; Synonym=Klaa0175, PK38;  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;  
 OC Muridae; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE [MRNA], AND TISSUE SPECIFICITY.  
 RC TISSUE=teratocyst;  
 RX MEDLINE=97449306; PubMed=9305775; DOI=10.1016/S0378-1119(97)00181-9;  
 RA Gil M., Yang Y., Lee Y., Choi I., Ha H.;  
 RT "Cloning and expression of a cDNA encoding a novel protein  
 RT serine/threonine kinase predominantly expressed in hematopoietic  
 RT cells.";  
 RL Gene 195:295-301(1997).  
 RN [2]  
 RP NUCLEOTIDE SEQUENCE [MRNA], TISSUE SPECIFICITY, AND DEVELOPMENTAL  
 RP STAGE.  
 RC STRAIN=C57BL/6 X DBA/2;  
 RX PubMed=9136115;  
 RX DOI=10.1002/(SICI)1098-2795(199706)47:2<148::AID-MRD4>3.0.CO;2-M;  
 RA Oyama R., Kasukawa T., Katayama S., Gough J., Frith M.C., Maeda N.,  
 RA Yamada R., Ravasi T., Lenhard B., Wells C., Kodzias R., Shimokawa K.,  
 RA Bajic V.B., Brenner S.E., Batalov S., Forrest A.R., Zavolan M.,  
 RA Davis M.J., Wilming L.G., Adkins J., Allen J.E.,  
 RA Ambesi-Impombato A., Apweiler R., Aturaliya R.N., Bailey T.L.,  
 RA Bansal M., Baxter L.W., Beisel K.W., Bersano T., Bono H., Chalk A.M.,  
 RA Chiu K.P., Choudhury V., Christoffels A., Clutterbuck D.R.,  
 RA Crowe M.L., Dalla E., Dall'Oy M., de Bono B., Della Gatta G.,  
 RA di Bernardo D., Down T., Engstrom P., Fagioli M., Faulkner G.,  
 RA Fletcher C.F., Fukushima T., Furuno M., Futaki S., Gariboldi M.,  
 RA Georgii-Hemming P., Gingeras T.R., Gojobori T., Green R.E.,  
 RA Gustigich S., Harbers M., Hayashi Y., Hensch T.K., Hirokawa N.,  
 RA Hill D., Hummelich L., Iacono M., Ikeo K., Iwama A., Ishikawa T.,  
 RA Jakt M., Kanapin A., Katoh M., Kawasawa Y., Kelso J., Kitamura H.,  
 RA Kitano H., Kollias G., Krishnan S.P., Kruger A., Kummerfeld S.K.,  
 RA Kurochkin I.V., Lareau L.P., Lazarevic D., Lipovich L., Liu J.,  
 RA Liuni S., McWilliam S., Madan Babu M., Madera M., Marchionni L.,  
 RA Matsuda H., Matsuzawa S., Miki H., Mignone F., Miyake S., Morris K.,  
 RA Mottagui-Tabar S., Mulder N., Nakano N., Nakaochi H., Ng P.,  
 RA Nilsson R., Nishiguchi S., Nishikawa S., Nori F., Ohara O.,  
 RA Okazaki Y., Orlando V., Pang K.C., Pavan W.J., Pavoni G., Pesole G.,  
 RA Petrovsky N., Piazza S., Reed J., Reid J.F., Ring B.Z., Ringwald M.,  
 RA Rost B., Ruan Y., Salzberg S.L., Sandelin A., Schneider C., Sheng Y.,  
 RA Schonbach C., Sekiguchi K., Semple C.A., Seno S., Sessa L., Sheng Y.,  
 RA Shibata Y., Shimada H., Shimada K., Silva D., Sinclair B.,  
 RA Sperling S., Stupka E., Sugiyama K., Sultana R., Takenaka Y., Taki K.,  
 RA Tammoja K., Tan S.L., Tang S., Taylor M.S., Tegner J., Teichmann S.A.,  
 RA Ueda H.R., van Nimwegen E., Verardo R., Wei C.L., Yagi K.,  
 RA Yamanishi H., Zabarovsky E., Zhu S., Zimmer A., Hide W., Bult C.,  
 RA Grimmond S.M., Tesada R.D., Liu E.T., Brusio V., Quackenbush J.,  
 RA Wahlestedt C., Mattick J.S., Hume D.A., Kai C., Sasaki D., Tomaru Y.,  
 RA Fukuda S., Kanamori-Katayama M., Suzuki M., Aoki J., Arakawa T.,  
 RA Iida J., Imamura K., Itoh M., Kato T., Kawaji H., Kawagashira N.,  
 RA Kawahara T., Kojima M., Kondo S., Konno H., Nakano K., Ninomiya N.,  
 RA Nishio T., Okada M., Plessey A., Shibata K., Shiraki T., Suzuki S.,  
 RA Tsgami T., Waki K., Watanabe A., Okamura-Oho Y., Suzuki H., Kawai J.,  
 RA Hayashizaki Y.;  
 RT "The transcriptional landscape of the mammalian genome."

RL Science 309:1559-1563(2005).  
 RN [4]  
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].  
 RC TISSUE=Embryonic tail;  
 RX MEDLINE=22977043; PubMed=14621295; DOI=10.1093/dnares/10.4.167;  
 RA Okazaki N., Kikuno R., Ohara R., Inamoto S., Koseki H., Hiraoka S.,  
 RA Saga Y., Nagase T., Ohara O., Koga H.;  
 RT "Prediction of the coding sequences of mouse homologues of KIAA gene:  
 RT III. The complete nucleotide sequences of 500 mouse KIAA-homologous  
 RT cDNAs identified by screening of terminal sequences of cDNA clones  
 RT randomly sampled from size-fractionated libraries.";  
 RL DNA Res. 10:167-180(2003).  
 CC -!- FUNCTION: Phosphorylates ZNF622 and may contribute to its  
 CC redirection to the nucleus. May be involved in the inhibition of  
 CC spliceosome assembly during mitosis (By similarity).  
 CC -!- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.  
 CC -!- SUBUNIT: Interacts with ZNF622 and PPIR8 (By similarity).  
 CC -!- SUBCELLULAR LOCATION: Cytoplasm (Potential).  
 CC -!- TISSUE SPECIFICITY: Expressed in testis, ovary, thymus, spleen and  
 CC T-cell.  
 CC -!- DEVELOPMENTAL STAGE: Expressed in the 2-cell-stage embryo,  
 CC followed by a strong expression at 8-cell-stage.  
 CC -!- PTM: Autophosphorylated (By similarity).  
 CC -!- SIMILARITY: Belongs to the Ser/Thr protein kinase family. SNF1  
 CC subfamily.  
 CC -!- SIMILARITY: Contains 1 KAI (kinase-associated) domain.  
 CC -----  
 CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms  
 CC Distributed under the Creative Commons Attribution-NoDerivs License  
 CC -----  
 DR EMBL: L76158; AAB72030.1; -; mRNA.  
 DR EMBL: X95351; AAG64641.1; -; mRNA.  
 DR EMBL: AK011932; BAB27923.1; -; mRNA.  
 DR EMBL: AK129076; BAC97886.1; -; mRNA.  
 DR HSSP: Q63450; 1A06.  
 DR MGI: MGI:106924; Melk.  
 DR GO: GO:0005737; Cytoplasm; IDA.  
 DR GO: GO:004672; F:protein kinase activity; IDA.  
 DR InterPro: IPR001772; Kinase C.  
 DR InterPro: IPR000719; Prot\_kinase.  
 DR InterPro: IPR008271; Ser\_thr\_pkin\_AS.  
 DR InterPro: IPR002290; Ser\_thr\_pkinase.  
 DR Pfam: PF02149; KAI; 1.  
 DR Pfam: PF00069; Pkinase; 1.  
 DR ProDom: PD000001; Prot\_kinase; 1.  
 DR SMART: SM00220; S\_TKc; 1.  
 DR PROSITE: PS00107; PROTEIN\_KINASE\_ATP; 1.  
 DR PROSITE: PS00111; PROTEIN\_KINASE\_DOM; 1.  
 DR PROSITE: PS00108; PROTEIN\_KINASE\_ST; 1.  
 KW ATP-binding; Kinase; Nucleotide-Binding; Phosphorylation;  
 KW Serine/threonine-protein kinase; Transferase.  
 FT CHAIN 1 643 Maternal embryonic leucine zipper kinase.  
 FT /FTID=PRO\_0000086324.  
 FT DOMAIN 11 263 Protein kinase.  
 FT DOMAIN 594 643 KAI.  
 FT NP\_BIND 17 25 ATP (By similarity).  
 FT ACT\_SITE 132 132 Proton acceptor (By similarity).  
 FT BINDING 40 40 ATP (By similarity).  
 FT CONFLICT 335 335 P -> L (in Ref. 3).  
 SQ SEQUENCE 643 AA; 72713 MW; 411792A8A19FE213 CRC64;  
 Query Match 83.1%; Score 2864; DB 1; Length 643;  
 Best Local Similarity 83.3%; Pred. No. 1.4e-184;  
 Matches 542; Conservative 45; Mismatches 56; Indels 8; Gaps 2;  
 QY 1 MKDYDELLKYELHETIGTGGFAKVKLACHILITGEMVAIKIMDKNTLGSGLPRIKTEIA 60  
 Db 1 MKDYDELLKYELHETIGTGGFAKVKLACHILITGEMVAIKIMDKNTLGSGLPRIKTEIA 60  
 QY 61 LKNLRHQHICQLYHVLETKNKIFMWLEPCPGGELFDYIISQRLSEETRVVPRQIVSAV 120  
 Db 61 LKSLRHHQICQLYHVLETKNKIFMWLEPCPGGELFDYIISQRLSEETRVVPRQIVSAV 120

QY 121 AYVHSQGYAHRDLKPEVHKLIDFGLCAKPKGNKDYHLQTCSSLAYAAPALI 180  
DB 121 AYVHSQGYAHRDLKPEVHKLIDFGLCAKPKGNKDYHLQTCSSLAYAAPALI 180  
QY 181 QGKSYLGSEADVMSGILLVLMCGFLPFDDNNMALKYKIMRGKYDVPKWLSPSSLTLL 240  
DB 181 QGKSYLGSEADVMSGILLVLMCGFLPFDDNNMALKYKIMRGKYDVPKWLSPSSLTLL 240  
QY 241 QOMLQVDPKCRISMRNLNHPWIMQDYNYPVQSKNPFILHDDDCVTELSVHRNNRQT 300  
DB 241 QOMLQVDPKCRISMRNLNHPWIMQDYNYPVQSKNPFILHDDDCVTELSVHRNNRQT 300  
QY 301 MEDLISLWQYDHTATYVLLLLAKARGKPVRLBSFSCQQAATPTDIKSNWNSLEDV 360  
DB 301 MEDLISLWQYDHTATYVLLLLAKARGKPVRLBSFSCQQAATPTDIKSNWNSLEDV 360  
QY 361 TASDKNVVAGLIDYDWCEDDLSTGAATPRTSQFTKYWTESGVESKSLTPALCRTLANKL 420  
DB 361 TASDKNVVAGLIDYDWCEDDLSTGAATPRTSQFTKYWTESGVESKSLTPALCRTLANKL 420  
QY 421 KXKNENYVTPKSAVKNBEYFMFPFKTPVKNQKHREILTPNRYTTPSKARNQCLKETPI 480  
DB 421 KXKNENYVTPKSAVKNBEYFMFPFKTPVKNQKHREILTPNRYTTPSKARNQCLKETPI 480  
QY 481 KIPVNSTGTDKMTGVISPPRRCSRVELDNOAHMBETPKRGAQVFGSLERGLDKVITV 540  
DB 481 KIPVNSTGTDKMTGVISPPRRCSRVELDNOAHMBETPKRGAQVFGSLERGLDKVITV 540  
QY 473 RTPGNSAGADTLTGTGVSPPRRCSRVDVNLQAHMEDTPKKGNTNVPFSGSLERGLDKVITA 532  
DB 473 RTPGNSAGADTLTGTGVSPPRRCSRVDVNLQAHMEDTPKKGNTNVPFSGSLERGLDKVITA 532  
QY 541 LTRSKRKGSGARDGPRRLKHLHYNTVTRLVNPDOLINEIMSLPKKHVDVFKGYTLKQOT 600  
DB 541 LTRSKRKGSGARDGPRRLKHLHYNTVTRLVNPDOLINEIMSLPKKHVDVFKGYTLKQOT 600  
QY 593 QSDFGKVTWQFEVLCQKPDVVGIRROKRGDAMVYKRLVEDILSSCKV 651  
DB 593 QSDFGKVTWQFEVLCQKPDVVGIRROKRGDAMVYKRLVEDILSSCKV 651  
RESULT 6  
ID Q3TJF3 MOUSE PRELIMINARY; PRT; 643 AA.  
AC Q3TJF3  
DT 11-OCT-2005, integrated into UniProtKB/TrEMBL.  
DT 07-FEB-2006, entry version 1.  
DE 12 days pregnant adult female placenta cDNA, RIKEN full-length  
DE enriched library, clone:1530014115 product:maternal embryonic leucine  
DE zipper kinase, full insert sequence.  
GN Name=Mek;  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;  
OC Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
[1]  
RN NUCLEOTIDE SEQUENCE.  
RC STRAIN=C57BL/6J; TISSUE=Placenta;  
RX PubMed=16141072; DOI=10.1126/science.1112014;  
RA Carninci P., Katayama T., Katayama S., Gough J., Frith M.C., Maeda N.,  
RA Oyama R., Ravasi T., Lenhard B., Wells C., Kodzius R., Shimokawa K.,  
RA Bajic V.B., Brenner S.E., Batalov S., Forrest A.R., Zavolan M.,  
RA Davis M.J., Wilming L.G., Aidinis V., Allen J.E.,  
RA Ambesi-Impombato A., Apweiler R., Aturaliya R.N., Bailey T.L.,  
RA Bansal M., Baxter L., Beisel K.W., Bersano T., Bono H., Chalk A.M.,  
RA Chiu K.P., Choudhary V., Christoffels A., Clutterbuck D.R.,  
RA Crowe M.L., Dalla E., Dalrymple B.P., de Bono B., Della Gatta G.,  
di Bernardo D., Down T., Engstrom P., Fagiolini M., Faulkner G.,  
Fletcher C.P., Fukushima T., Furuno M., Futaki S., Gariboldi M.,  
Georgii-Hemming P., Gingeras T.R., Gojobori T., Green R.E.,  
Gustincich S., Hayber M., Hayashi Y., Hensch T.K., Hirokawa N.,  
Hill D., Huminec L., Iacono M., Ikeo K., Iwama A., Ishikawa T.,  
Jakt M., Kanapin A., Katoh M., Kawasawa Y., Kelso J., Kitamura H.,  
Kitano H., Kollias G., Krishnan S.P., Kruger A., Kummerfeld S.K.,  
Kurochkin I.V., Lareau L.F., Lazarevic D., Lipovich L., Liu J.,  
Liu S., McWilliam S., Madan Babu M., Madera M., Marchionni L.,  
Matsuda H., Matsuzawa S., Miki H., Mignone F., Miyake S., Morris K.,  
Motteguel-Tabar S., Mulder N., Nakano N., Nakachi H., Ng P.,  
Nilsson R., Nishiguchi S., Nishikawa S., Nori F., Ohara O.,  
Okazaki Y., Orlando V., Pang K.C., Pavan W.J., Pavese G., Pesole G.,  
Petkovsky N., Piazza S., Reed J., Reid J.F., Ring B.Z., Ringwald M.,  
Rost B., Ruan Y., Salzberg S.L., Sandelin A., Schneider C.,  
Schonbach C., Sekiguchi K., Sempole C.A., Sene S., Sessa L., Sheng Y.,  
Shibata Y., Shimada H., Shimada K., Silva D., Sinclair B.,  
Sperling S., Stupka E., Sugura K., Sultana R., Takenaka Y., Taki K.,  
Tammaja K., Tan S.L., Tang S., Taylor M.S., Tegner J., Teichmann S.A.,  
Ueda H.R., van Nimwegen E., Verardo R., Wei C.L., Yagi K.,  
Yamanishi H., Zabarovsky E., Zhu S., Zimmer A., Hide W., Bult C.,  
Grimmond S.M., Teasdale R.D., Liu E.T., Brusic V., Quackenbush J.,  
Wahlstedt C., Mattick J.S., Hume D.A., Kai C., Sasaki D., Tomaru Y.,  
Fukuda S., Kanamori-Katayama M., Suzuki M., Aoki J., Arakawa T.,  
Iida J., Inamura K., Itoh M., Kato T., Kawaji H., Kawagashira N.,  
Kawashima T., Kojima M., Kondo S., Konno H., Nakano K., Nanomiya N.,  
Nishio T., Okada M., Plessy C., Shibata K., Shiraki T., Suzuki S.,  
Tagami M., Waki K., Watahiki A., Okamura-Oho Y., Suzuki H., Kawai J.,  
Hayaishizaki Y.  
RT "The transcriptional landscape of the mammalian genome."  
RL Science 309:1559-1563 (2005).  
RN [3]  
RP NUCLEOTIDE SEQUENCE.  
RC STRAIN=C57BL/6J; TISSUE=Placenta;  
RX PubMed=16141073; DOI=10.1126/science.1112009;  
RG RIKEN Genome Exploration Research Group, and Genome Science Group  
RG (Genome Network Core Team) and the FANTOM Consortium;  
RT "Antisense Transcription in the Mammalian Transcriptome."  
RL Science 309:1564-1566 (2005).  
RN [4]  
RP NUCLEOTIDE SEQUENCE.  
RC STRAIN=C57BL/6J; TISSUE=Placenta;  
RX PubMed=12466851; DOI=10.1038/nature01266;  
RA Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,  
RA Nakaide I., Osato N., Saito R., Suzuki H., Yamanaka I., Kiyosawa H.,  
RA Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojobori T.,  
RA Balderelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,  
RA Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,  
RA Blake J.A., Bradt D., Brusic V., Chothia C., Corbani L.E., Cousins S.,  
RA Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,  
RA Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,  
RA Grimmond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,  
RA Kanai A., Kawaji H., Kawasawa Y., Kedzierski R.M., King B.L.,  
RA Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,  
RA Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,  
RA Nagashima T., Numa K., Okido T., Pavan W.J., Pertea G., Pesole G.,  
RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,  
RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,  
RA Sandelin A., Schneider C., Sempole C.A., Setou M., Shimada K.,  
RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,  
RA Verardo R., Wagner L., Wahlstedt C., Wang Y., Watanabe Y., Wells C.,  
RA Wilming L.G., Wynshaw-Boris A., Yanagisawa M., Yang I., Yang L.,  
RA Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,  
RA Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,  
RA Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,  
RA Hara A., Hashizume W., Imotani K., Iehli Y., Itoh M., Kagawa I.,  
RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,  
RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,  
Birney E., Hayaishizaki Y.;  
RT "Analysis of the mouse transcriptome based on functional annotation of  
RT 60,770 full-length cDNAs."  
RL Nature 420:563-573 (2002).  
RN [5]

RA 121 AYVHSQGYAHRDLKPEVHKLIDFGLCAKPKGNKDYHLQTCSSLAYAAPALI 180  
RA 121 AYVHSQGYAHRDLKPEVHKLIDFGLCAKPKGNKDYHLQTCSSLAYAAPALI 180  
RA 181 QGKSYLGSEADVMSGILLVLMCGFLPFDDNNMALKYKIMRGKYDVPKWLSPSSLTLL 240  
RA 181 QGKSYLGSEADVMSGILLVLMCGFLPFDDNNMALKYKIMRGKYDVPKWLSPSSLTLL 240  
RA 241 QOMLQVDPKCRISMRNLNHPWIMQDYNYPVQSKNPFILHDDDCVTELSVHRNNRQT 300  
RA 241 QOMLQVDPKCRISMRNLNHPWIMQDYNYPVQSKNPFILHDDDCVTELSVHRNNRQT 300  
RA 301 MEDLISLWQYDHTATYVLLLLAKARGKPVRLBSFSCQQAATPTDIKSNWNSLEDV 360  
RA 301 MEDLISLWQYDHTATYVLLLLAKARGKPVRLBSFSCQQAATPTDIKSNWNSLEDV 360  
RA 361 TASDKNVVAGLIDYDWCEDDLSTGAATPRTSQFTKYWTESGVESKSLTPALCRTLANKL 420  
RA 361 TASDKNVVAGLIDYDWCEDDLSTGAATPRTSQFTKYWTESGVESKSLTPALCRTLANKL 420  
RA 421 KXKNENYVTPKSAVKNBEYFMFPFKTPVKNQKHREILTPNRYTTPSKARNQCLKETPI 480  
RA 421 KXKNENYVTPKSAVKNBEYFMFPFKTPVKNQKHREILTPNRYTTPSKARNQCLKETPI 480  
RA 481 KIPVNSTGTDKMTGVISPPRRCSRVELDNOAHMBETPKRGAQVFGSLERGLDKVITV 540  
RA 481 KIPVNSTGTDKMTGVISPPRRCSRVELDNOAHMBETPKRGAQVFGSLERGLDKVITV 540  
RA 473 RTPGNSAGADTLTGTGVSPPRRCSRVDVNLQAHMEDTPKKGNTNVPFSGSLERGLDKVITA 532  
RA 473 RTPGNSAGADTLTGTGVSPPRRCSRVDVNLQAHMEDTPKKGNTNVPFSGSLERGLDKVITA 532  
RA 541 LTRSKRKGSGARDGPRRLKHLHYNTVTRLVNPDOLINEIMSLPKKHVDVFKGYTLKQOT 600  
RA 541 LTRSKRKGSGARDGPRRLKHLHYNTVTRLVNPDOLINEIMSLPKKHVDVFKGYTLKQOT 600  
RA 593 QSDFGKVTWQFEVLCQKPDVVGIRROKRGDAMVYKRLVEDILSSCKV 651  
RA 593 QSDFGKVTWQFEVLCQKPDVVGIRROKRGDAMVYKRLVEDILSSCKV 651  
RESULT 6  
ID Q3TJF3 MOUSE PRELIMINARY; PRT; 643 AA.  
AC Q3TJF3  
DT 11-OCT-2005, integrated into UniProtKB/TrEMBL.  
DT 07-FEB-2006, entry version 1.  
DE 12 days pregnant adult female placenta cDNA, RIKEN full-length  
DE enriched library, clone:1530014115 product:maternal embryonic leucine  
DE zipper kinase, full insert sequence.  
GN Name=Mek;  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;  
OC Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
[1]  
RN NUCLEOTIDE SEQUENCE.  
RC STRAIN=C57BL/6J; TISSUE=Placenta;  
RX PubMed=16141072; DOI=10.1126/science.1112014;  
RA Carninci P., Katayama T., Katayama S., Gough J., Frith M.C., Maeda N.,  
RA Oyama R., Ravasi T., Lenhard B., Wells C., Kodzius R., Shimokawa K.,  
RA Bajic V.B., Brenner S.E., Batalov S., Forrest A.R., Zavolan M.,  
RA Davis M.J., Wilming L.G., Aidinis V., Allen J.E.,  
RA Ambesi-Impombato A., Apweiler R., Aturaliya R.N., Bailey T.L.,  
RA Bansal M., Baxter L., Beisel K.W., Bersano T., Bono H., Chalk A.M.,  
RA Chiu K.P., Choudhary V., Christoffels A., Clutterbuck D.R.,  
RA Crowe M.L., Dalla E., Dalrymple B.P., de Bono B., Della Gatta G.,  
di Bernardo D., Down T., Engstrom P., Fagiolini M., Faulkner G.,  
Fletcher C.P., Fukushima T., Furuno M., Futaki S., Gariboldi M.,  
Georgii-Hemming P., Gingeras T.R., Gojobori T., Green R.E.,  
Gustincich S., Hayber M., Hayashi Y., Hensch T.K., Hirokawa N.,  
Hill D., Huminec L., Iacono M., Ikeo K., Iwama A., Ishikawa T.,  
Jakt M., Kanapin A., Katoh M., Kawasawa Y., Kelso J., Kitamura H.,  
Kitano H., Kollias G., Krishnan S.P., Kruger A., Kummerfeld S.K.,  
Kurochkin I.V., Lareau L.F., Lazarevic D., Lipovich L., Liu J.,  
Liu S., McWilliam S., Madan Babu M., Madera M., Marchionni L.,  
Matsuda H., Matsuzawa S., Miki H., Mignone F., Miyake S., Morris K.,  
Motteguel-Tabar S., Mulder N., Nakano N., Nakachi H., Ng P.,  
Nilsson R., Nishiguchi S., Nishikawa S., Nori F., Ohara O.,  
Okazaki Y., Orlando V., Pang K.C., Pavan W.J., Pavese G., Pesole G.,  
Petkovsky N., Piazza S., Reed J., Reid J.F., Ring B.Z., Ringwald M.,  
Rost B., Ruan Y., Salzberg S.L., Sandelin A., Schneider C.,  
Schonbach C., Sekiguchi K., Sempole C.A., Sene S., Sessa L., Sheng Y.,  
Shibata Y., Shimada H., Shimada K., Silva D., Sinclair B.,  
Sperling S., Stupka E., Sugura K., Sultana R., Takenaka Y., Taki K.,  
Tammaja K., Tan S.L., Tang S., Taylor M.S., Tegner J., Teichmann S.A.,  
Ueda H.R., van Nimwegen E., Verardo R., Wei C.L., Yagi K.,  
Yamanishi H., Zabarovsky E., Zhu S., Zimmer A., Hide W., Bult C.,  
Grimmond S.M., Teasdale R.D., Liu E.T., Brusic V., Quackenbush J.,  
Wahlstedt C., Mattick J.S., Hume D.A., Kai C., Sasaki D., Tomaru Y.,  
Fukuda S., Kanamori-Katayama M., Suzuki M., Aoki J., Arakawa T.,  
Iida J., Inamura K., Itoh M., Kato T., Kawaji H., Kawagashira N.,  
Kawashima T., Kojima M., Kondo S., Konno H., Nakano K., Nanomiya N.,  
Nishio T., Okada M., Plessy C., Shibata K., Shiraki T., Suzuki S.,  
Tagami M., Waki K., Watahiki A., Okamura-Oho Y., Suzuki H., Kawai J.,  
Hayaishizaki Y.  
RT "The transcriptional landscape of the mammalian genome."  
RL Science 309:1559-1563 (2005).  
RN [3]  
RP NUCLEOTIDE SEQUENCE.  
RC STRAIN=C57BL/6J; TISSUE=Placenta;  
RX PubMed=16141073; DOI=10.1126/science.1112009;  
RG RIKEN Genome Exploration Research Group, and Genome Science Group  
RG (Genome Network Core Team) and the FANTOM Consortium;  
RT "Antisense Transcription in the Mammalian Transcriptome."  
RL Science 309:1564-1566 (2005).  
RN [4]  
RP NUCLEOTIDE SEQUENCE.  
RC STRAIN=C57BL/6J; TISSUE=Placenta;  
RX PubMed=12466851; DOI=10.1038/nature01266;  
RA Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,  
RA Nakaide I., Osato N., Saito R., Suzuki H., Yamanaka I., Kiyosawa H.,  
RA Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojobori T.,  
RA Balderelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,  
RA Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,  
RA Blake J.A., Bradt D., Brusic V., Chothia C., Corbani L.E., Cousins S.,  
RA Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,  
RA Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,  
RA Grimmond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,  
RA Kanai A., Kawaji H., Kawasawa Y., Kedzierski R.M., King B.L.,  
RA Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,  
RA Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,  
RA Nagashima T., Numa K., Okido T., Pavan W.J., Pertea G., Pesole G.,  
RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,  
RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,  
RA Sandelin A., Schneider C., Sempole C.A., Setou M., Shimada K.,  
RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,  
RA Verardo R., Wagner L., Wahlstedt C., Wang Y., Watanabe Y., Wells C.,  
RA Wilming L.G., Wynshaw-Boris A., Yanagisawa M., Yang I., Yang L.,  
RA Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,  
RA Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,  
RA Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,  
RA Hara A., Hashizume W., Imotani K., Iehli Y., Itoh M., Kagawa I.,  
RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,  
RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,  
Birney E., Hayaishizaki Y.;  
RT "Analysis of the mouse transcriptome based on functional annotation of  
RT 60,770 full-length cDNAs."  
RL Nature 420:563-573 (2002).  
RN [5]

RP NUCLEOTIDE SEQUENCE.  
RC STRAIN=C57BL/6J; TISSUE=Placenta; DOI=10.1038/35055500;  
RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;  
RA Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,  
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,  
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,  
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,  
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,  
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,  
RA Kuehl P., Lewis S., Matsuo Y., Nikaïdo I., Peele G., Quackenbush J.,  
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,  
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,  
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.P.,  
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,  
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,  
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,  
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,  
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,  
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,  
RA Wynehaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,  
RA Hayashizaki Y.;  
RT "Functional annotation of a full-length mouse cDNA collection.";  
RL Nature 409:685-690(2001).  
RN [6]

RP NUCLEOTIDE SEQUENCE.  
RC STRAIN=C57BL/6J; TISSUE=Placenta;  
RX MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;  
RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,  
RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;  
RT "Normalization and subtraction of cap-trapper-selected cDNAs to  
prepare full-length cDNA libraries for rapid discovery of new genes.";  
RL Genome Res. 10:1617-1630(2000).  
RN [7]

RP NUCLEOTIDE SEQUENCE.  
RC STRAIN=C57BL/6J; TISSUE=Placenta;  
RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;  
RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,  
RA Konno H., Akiyama J., Nishi K., Kiteunai T., Tashiro H., Itoh M.,  
RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,  
RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,  
RA Fujiwaka S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,  
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuura S., Kawai J.,  
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;  
RT "RIKEN integrated sequence analysis (RISA) system-384-format  
sequencing pipeline with 384 multicapillary sequencer.";  
RL Genome Res. 10:1757-1771(2000).  
RN [8]

RP NUCLEOTIDE SEQUENCE.  
RC STRAIN=C57BL/6J; TISSUE=Placenta;  
RA Arakawa T., Carninci P., Fukuda S., Haghizume W., Hayaahida K.,  
RA Hori F., Iida J., Imamura K., Imotani K., Itoh M., Kanagawa S.,  
RA Kawai J., Kojima M., Konno H., Murata M., Nakamura M., Ninomiya N.,  
RA Nishiyori H., Nomura K., Ohno M., Sakazume N., Sano H., Sasaki D.,  
RA Shibata K., Shiraki T., Tagami M., Tagami Y., Waki H., Watahiki A.,  
RA Muramatsu M., Hayashizaki Y.;  
RL Submitted (APR-2004) to the EMBL/GenBank/DBJ databases  
CC -1- SIMILARITY: Belongs to the Ser/Thr protein kinase family.  
CC -----  
CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>  
CC Distributed under the Creative Commons Attribution-NoDerivs License  
CC -----  
CC EMBL: AK167457; BAE39542.1; -; mRNA.  
CC MGI: MGI:108924; Melk.  
CC GO: GO:0005737; C:cytoplasm; IDA.  
CC GO: GO:0004713; F:protein-tyrosine kinase activity; RCA.  
CC InterPro: IPR001772; Kinase C.  
CC InterPro: IPR000719; Prot\_kinase.  
CC InterPro: IPR008271; Ser\_thr\_pkin\_AS.  
CC InterPro: IPR002290; Ser\_thr\_pkinase.  
CC InterPro: IPR001245; Tyr\_pkinase.  
CC Pfam: PF02149; KAI; 1.  
CC Pfam: PF00069; Pkinase; 1.  
CC ProDom: PD000001; Prot\_kinase; 1.

DR SMART; SM00220; S\_TKc; 1.  
DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; 1.  
DR PROSITE; PS00111; PROTEIN\_KINASE\_DOM; 1.  
DR PROSITE; PS00108; PROTEIN\_KINASE\_ST; UNKNOWN\_1.  
KW ATP-binding; Kinase; Nucleotide-binding;  
KW Serine/threonine-protein kinase; transferase;  
SQ SEQUENCE 643 AA; 72757 MW; ADA9240C6C56FA4 CRC64;  
Query Match 83.1%; Score 2863; DB 2; Length 643;  
Best Local Similarity 83.3%; Pred. No. 1.6e-184;  
Matches 542; Conservative 45; Mismatches 56; Indels 8; Gaps 2;  
QY 1 MKDYDELLKYYELHETHTGTGGSPAKVLACHILITGEMVAIKIMDKNTLGSDDLPRITKEISA 60  
DB 1 MKDYDELLKYYELHETHTGTGGSPAKVLACHILITGEMVAIKIMDKNTLGSDDLPRITKEISA 60  
QY 61 LKVLRHQHICQLYHVLETKANKIFMWLEVCPCGELFDYIISODRLSEETRVVPRQIVSAV 120  
DB 61 LKSLRHQHICQLYHVLETKANKIFMWLEVCPCGELFDYIISODRLSEETRVVPRQIVSAV 120  
QY 121 AVVHSQGYAHRDLKPENLFFDEYHKLIDFGLCAKPKGNKDYHLQTCGSLAYAAPELI 180  
DB 121 AVVHSQGYAHRDLKPENLFFDENHKLIDFGLCAKPKGNKDYHLQTCGSLAYAAPELI 180  
QY 181 QGKSYLGSEADYVSMGILLVYLMCGFLPFDDNNVMAIYKKIMRGKYDVPKWLSPSSILL 240  
DB 181 QGKSYLGSEADYVSMGILLVYLMCGFLPFDDNNVMAIYKKIMRGKYDVPKWLSPSSILL 240  
QY 241 QOMLQVDPKRIKSMKNLLNHPIMQDYNYPVWQSKNPIHLDDCVTSLVHHRRNQOT 300  
DB 241 QOMLQVDPKRIKSMKNLLNHPIMQDYNYPVWQSKNPIHLDDCVTSLVHHRRNQOT 300  
QY 301 MEDLISLMOYDHLTATYLLLLAKKARGKPVRLRLSSFCQASATPFTDIKSNNSLEDV 360  
DB 301 MEDLISLMOYDHLTATYLLLLAKKARGKPVRLRLSSFCQASATPFTDIKSNNSLEDV 360  
QY 361 TASDKNYVAGLIDYDWCEDDLSTGAATPRTSFTKYWTESNGVESKSLTPALCRTPANKL 420  
DB 357 STSDNVCNAGLIDYELCEDKL---LAPKTPQVTKHLAESAHAASPKSPAGVRRVAVANKL 412  
QY 421 KKNENYTPKSAVQVEEYFMPPEKTPYQNKQHKREILLTPNRYTPPSKARNOCLKETPI 480  
DB 413 MDKENVCTPKSVQNEEQVFSEPKIPVSKNKKYKEIPASPTFRFPFPAKARACQLEAPV 472  
QY 481 KIPVNSTGDKLMTGVISPERCSRVELDLNOAHMEETPKRGAKVFGSLERGLDKVITV 540  
DB 473 RFPGNSAGADTLTTGVISPERCSRSDVDLNOAHMEDTPKKGTVNFGSLERGLDKVITA 532  
QY 541 LTRSKRKGSGARDGPRRLKLHYNVTTTLVNPOLLNEIMSILPKKHVDVQKGYTLKCOQ 600  
DB 533 LTRNKKGSGARDGPRKRLHYNVTTTLVNPOLLSEIMAILPKKNVDFVQKGYTLKCOQ 592  
QY 601 QSDFGKVTWQFLEVCQKQKPDVGIIRRLKGDWVYRELVDILSSCKV 651  
DB 593 QSDFGKVTWQFLEVCQKQKQRPDVGIIRRLKGDWVYRELVDILSSCKM 643

RESULT 7  
Q3TPUL\_MOUSE  
ID Q3TPUL\_MOUSE PRELIMINARY; PRT; 643 AA.  
AC Q3TPUL;  
DT 11-OCT-2005, integrated into UniProtKB/TrEMBL.  
DT 11-OCT-2005, sequence version 1.  
DT 07-FEB-2006, entry version 7.  
DE 12 days embryo spinal cord cDNA, RIKEN full-length enriched library,  
DE clone: C530047P3 product: maternal embryonic leucine zipper kinase,  
DE full insert sequence (4 cells embryo 4 cells cDNA, RIKEN full-length  
DE enriched library, clone: I0C001106 product: maternal embryonic leucine  
DE zipper kinase, full insert sequence).  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;  
OC Muroidae; Muridae; Murinae; Mus.

OX NCBI\_TaxID=10090;  
 RN [1]  
 RN NUCLEOTIDE SEQUENCE.  
 RC STRAIN=C57BL/6J; TISSUE=Spinal cord;  
 RX MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;  
 RA Carninci P., Hayashizaki Y.;  
 RT "High-efficiency full-length cDNA cloning.";  
 RL Methods Enzymol. 303:19-44(1999).  
 RN [2]  
 RN NUCLEOTIDE SEQUENCE.  
 RC STRAIN=C57BL/6J; TISSUE=Spinal cord;  
 RX PubMed=16141072; DOI=10.1126/science.1112014;  
 RA Carninci P., Kasukawa T., Katayama S., Gough J., Frith M.C., Maeda N., Oyama R., Ravasi T., Lenhard B., Wells C., Kodzius B., Zvolan A., Shimokawa K., Bajic V.B., Brenner S.E., Batalov S., Forrest A.R., Zavolan M., Davis M.J., Wilming L.G., Adkins J.E., Allen J.E., Ambesi-Impombato A., Anweiler R., Aturaliya R.N., Bailey T.L., Bansal M., Baxter L., Beisel K.W., Bersano T., Bono H., Chalk A.M., Chiu K.P., Choudhary V., Christoffels A., Clutterbuck D.R., Crowe M.L., Dalla E., Dalrymple B.P., de Bono B., Della Gatta G., Di Bernardo D., Down T., Engstrom P., Fagioli M., Faulkner G., Fletcher C.F., Fukushima T., Furuno M., Futaki S., Gariboldi M., Georgii-Heminger P., Gingeras T.R., Gojobori T., Green R.E., Gustincich S., Harbers M., Hayashi Y., Hensch T.K., Hirokawa N., Hill D., Hummelbeck L., Iacono M., Ikeo K., Iwama A., Ishikawa T., Jakt M., Kanapin A., Katoh M., Kawasawa Y., Kelso J., Kitamura H., Kitano H., Kollias G., Krishnan S.P., Kruger A., Kummerfeld S.K., Kurochkin I.V., Lareau L.F., Lazarevic D., Lipovich L., Liu J., Liuni S., McWilliam S., Madan Babu M., Madera M., Marchionni L., Matsuda H., Matsuzawa S., Miki H., Mignone P., Miyake S., Morris K., Mottagui-Tabar S., Mulder N., Nakano N., Nakaguchi H., Ng P., Nilsson R., Nishiguchi S., Nishikawa S., Nori F., Ohara O., Okazaki Y., Orlando V., Pang K.C., Pavan W.J., Pavese G., Pesole G., Petrovsky N., Piazza S., Reed J.C., Reid J.P., Ring B.Z., Ringwald M., Roost B., Ruan Y., Salzberg S.L., Sandelin A., Schneider C., Schonbach C., Sekiguchi K., Sempile C.A., Seno S., Sessa L., Sheng Y., Shibata Y., Shimada H., Shimada K., Silva D., Sinclair B., Sperling S., Stupka E., Sugura K., Sultana R., Takenaka Y., Taki K., Tammoja K., Tan S.L., Tang S., Taylor M.S., Tegner J., Teichmann S.A., Ueda H.R., van Nimwegen E., Verardo R., Wei C.L., Yagi K., Yamanishi H., Zabarovsky E., Zhu S., Zimmer A., Hide W., Bult C., Grimmond S.M., Teasdale R.D., Liu E.T., Brusic V., Quackenbush J., Wahlestedt C., Mattick J.S., Hume D.A., Kai C., Sasaki D., Tomaru Y., Fukuda S., Kanamori-Katayama M., Suzuki M., Aoki J., Arakawa T., Iida J., Imamura K., Itoh M., Kato T., Kawaji H., Kawagashira N., Kawashima T., Kojima M., Kondo S., Konno H., Nakano K., Ninomiya N., Nishio T., Okada M., Pleasy C., Shibata K., Shiraki T., Suzuki S., Tagami M., Waki K., Watanabe A., Okamura-Oho Y., Suzuki H., Kawai J., Hayashizaki Y.;  
 RT "The transcriptional landscape of the mammalian genome.";  
 RL Science 309:1559-1563(2005).  
 RN [3]  
 RN NUCLEOTIDE SEQUENCE.  
 RC STRAIN=C57BL/6J; TISSUE=Spinal cord;  
 RX PubMed=16141073; DOI=10.1126/science.1112009;  
 RG RIKEN Genome Exploration Research Group, and Genome Science Group (Genome Network Core Team) and the FANTOM Consortium;  
 RT "Antisense Transcription in the Mammalian Transcriptome.";  
 RL Science 309:1564-1566(2005).  
 RN [4]  
 RN NUCLEOTIDE SEQUENCE.  
 RC STRAIN=C57BL/6J; TISSUE=Spinal cord;  
 RX MEDLINE=22354683; PubMed=12466851; DOI=10.1038/nature01266;  
 RA Nikaido I., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S., Yagi K., Tomaru Y., Ose T., Saito R., Suzuki H., Yamanaka I., Kiyosawa H., Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J., Schriml L.M., Katapin A., Matsuda H., Batalov S., Beisel K.W., Blake J.A., Bradt D., Brusic V., Chothia C., Corbani L.E., Cousins S., Dalla E., Dragani T.A., Fletcher C.P., Forrest A., Frazer K.S., Gaasterland T., Gariboldi M., Glissi C., Godzik A., Gough J., Grimmond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D., Kanai A., Kawaji H., Kawasawa Y., Kedzierski R.M., King B.L., Kanai A., Kawaji H., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A., Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H., Nagashima T., Numata K., Okido T., Pavan W.J., Perle G., Pesole G., Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S., Ravasi T., Reed J.C., Reid J.C., Reid J., Ring B.Z., Ringwald M., Sandelin A., Schneider C., Sempile C.A., Setou M., Shimada K., Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M., Ueda H.R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C., Wilming L.G., Wynshaw-Boris A., Yanagisawa M., Yang I., Yang L., Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayashizaki Y., Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K., Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S., Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I., Kawai J., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I., Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A., Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J., Birney E., Hayashizaki Y.;  
 RT "Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs.";  
 RL Nature 420:563-573(2002).  
 RN [5]  
 RN NUCLEOTIDE SEQUENCE.  
 RC STRAIN=C57BL/6J; TISSUE=Spinal cord;  
 RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;  
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y., Arakawa T., Hara A., Fukunishi Y., Kono H., Adachi J., Fukuda S., Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I., Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R., Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T., Fleschmann W., Gaasterland T., Glissi C., King B., Kochiwa H., Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J., Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T., Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G., Blake J., Moffell D., Bojunga N., Carninci P., de Bonaldo M.F., Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M., Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H., Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P., Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N., Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F., Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilming L., Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S., Hayashizaki Y.;  
 RT "Functional annotation of a full-length mouse cDNA collection.";  
 RL Nature 409:685-690(2001).  
 RN [6]  
 RN NUCLEOTIDE SEQUENCE.  
 RC STRAIN=C57BL/6J; TISSUE=Spinal cord;  
 RX MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;  
 RA Carninci P., Shibata Y., Hayatsu M., Sugahara Y., Shibata K., Itoh M., Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;  
 RT "Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes.";  
 RL Genome Res. 10:1617-1630(2000).  
 RN [7]  
 RN NUCLEOTIDE SEQUENCE.  
 RC STRAIN=C57BL/6J; TISSUE=Spinal cord;  
 RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;  
 RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P., Konno H., Akiyama J., Nishi K., Kitsuai T., Tashiro H., Itoh M., Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A., Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K., Fujiwara S., Inoue K., Togawa Y., Izawa M., Ohara E., Watanabe M., Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsura S., Kawai J., Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;  
 RT "RIKEN integrated sequence analysis (RISA) system-384-format sequencing pipeline with 384 multicapillary sequencer.";  
 RL Genome Res. 10:1757-1771(2000).  
 RN [8]  
 RN NUCLEOTIDE SEQUENCE.  
 RC STRAIN=C57BL/6J; TISSUE=Spinal cord;  
 RX Arakawa T., Carninci P., Fukuda S., Hashizume W., Hayashida K., Hori F., Iida J., Imamura K., Imotani K., Itoh M., Kanagawa S., Kawai J., Kojima M., Konno H., Murata M., Nakamura M., Ninomiya N., Nishiyori H., Nomura K., Ohno M., Sakazume N., Sano H., Sasaki D.,

```
RA Shibata K., Shiraki T., Tagami M., Tagami Y., Waki K., Watahiki A.,
RA Muramatsu M., Hayashizaki Y.;
RL Submitted (APR-2004) to the EMBL/GenBank/DBJ databases.
RW [9]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J;
RA Arakawa T., Carninci P., Fukuda S., Hashizume W., Hayashida K.,
RA Hori F., Iida J., Imamura K., Imotani K., Itoh M., Kanagawa S.,
RA Kawai J., Kojima M., Konno H., Murata M., Nakamura M., Ninomiya N.,
RA Nishiyori H., Nomura K., Ohno M., Sakazume N., Sano H., Sasaki D.,
RA Shibata K., Shiraki T., Tagami M., Tagami Y., Waki K., Watahiki A.,
RA Muramatsu M., Hayashizaki Y.;
RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.
CC 1- SIMILARITY: Belongs to the Ser/Thr protein kinase family.
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs license
CC -----
DR EMBL; AK164138; BAE37644.1; -; mRNA.
DR EMBL; AK145316; BAE26362.1; -; mRNA.
DR GO; GO:0005737; Cytoplasm; IDA.
DR GO; GO:0004713; F:protein-tyrosine kinase activity; RCA.
DR InterPro; IPR001772; Kinase_C.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR008271; Ser_thr_pkin_AS.
DR InterPro; IPR002290; Ser_thr_pkinase.
DR InterPro; IPR001245; Tyr_pkinase.
DR Pfam; PF02149; KAI; 1.
DR Pfam; PF00069; Kinase; 1.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00220; S_TKC; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00101; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; UNKNOWN_1.
KW ATP-binding; Kinase; Nucleotide-binding;
KW Serine/threonine-protein kinase; Transferase.
SQ SEQUENCE 643 AA; 72729 MW; 5637D2A4E8CFA216 CRC64;

Query Match 83.1%; Score 2863; DB 2; Length 643;
Best Local Similarity 93.3%; Pred. No. 1.6e-184;
Matches 542; Conservative 45; Mismatches 56; Indels 8; Gaps 2;

QY 1 MKDYDELLKYELHETIGTGGAFAKVLACHILTGEMVAIKIMDKNTGLSDLPRIKTEISA 60
DB 1 MKDYDELLKYELYETIGTGGAFAKVLACHVLTGEMVAIKIMDKNALGSDLPVKTETIDA 60

QY 61 LKNLRHQHICQLYHVLETANKIPWLEYCPGGBELFDYIISQDRLSEETRVVPRQIVSAV 120
DB 61 LKSLRHQHICQLYHVLETANKIPWLEYCPGGBELFDYIISQDRLSEETRVVPRQIVSAV 120

QY 121 AVVHSGYAHRLDKPENLLFDEYHKLIDFGLCAKPKGNKDYHLQTCGSLAYAAPELI 180
DB 121 AVVHSGYAHRLDKPENLLFDEYHKLIDFGLCAKPKGNKDYHLQTCGSLAYAAPELI 180

QY 181 QGKSYLGSEADVMSGILLYVLMCGFLPFDDDDNVMALYKIMRGKYDVPKWLSPSSILL 240
DB 181 QGKSYLGSEADVMSGILLYVLMCGFLPFDDDDNVMALYKIMRGKYDVPKWLSPSSILL 240

QY 241 QQMLQVDPKRTSMKNLLNHPWIMQDYNVPVEHQSNKPIHLDDDCVTELNVHHRNROT 300
DB 241 QQMLQVDPKRTSMRNLNHPWIMQDYSQVPEWQSKTPTLTHLDDCVTELNVHHRSSROT 300

QY 301 MEDLISLWQYDHLTATYLLALLAKARGKPVRLRLSFCGQASATPTFTDKSNWLSLEDV 360
DB 301 MEDLISLWQYDHLTATYLLALLAKARGKPARLQQLSFCGQASTPT-----KSNLSLEDM 356

QY 361 TASDKNXYAGLIDYDCCEDDLSTGAATPRTSQFTKYWTESNGVESKSLTPALCRTPANKL 420
DB 361 TASDKNXYAGLIDYDCCEDDLSTGAATPRTSQFTKYWTESNGVESKSLTPALCRTPANKL 420

QY 357 STSDDCNVCAGLIDYELCEDKL-----LAPKTPQVTKHLAESNHAASKSPAGVGRVAVANKL 412
DB 357 STSDDCNVCAGLIDYELCEDKL-----LAPKTPQVTKHLAESNHAASKSPAGVGRVAVANKL 412

QY 421 KKNENYTPKSAVKNSEYFMFPFKPTPVNKNQHKRILITTPRYTTPSKARNOCLKETPI 480
DB 421 KKNENYTPKSAVKNSEYFMFPFKPTPVNKNQHKRILITTPRYTTPSKARNOCLKETPI 480

QY 413 MDKENYCTPKSSVKNSEQVFSEPKIPVSKNQYKREIPASPTFRPTPAKARAQCLREAPV 472
DB 413 MDKENYCTPKSSVKNSEQVFSEPKIPVSKNQYKREIPASPTFRPTPAKARAQCLREAPV 472
```

```
QY 481 KIPVNSTGTDKLMTGVISPERCRSVELDLNOAHMEETPKRKGAQVFGSLERGLDKVITV 540
DB 473 RTPGNSAGADTLTTGTVISPERCRSMDVDLNOAHMEDTPKKGTNVFGSLERGLDKVLT 532

QY 541 LTRSRKRGSGARDGPRRLKLYHNVTTTTLVNPOLLNEMSLIPKXHVDFVQKGYTLKCOT 600
DB 533 LTRNFKKGSARDGPRKRLKLYHNVTTTTLVNPOLLSEIMAILPKKNVDFVQKGYTLKCOT 592

QY 601 QSDFGKVTMQPFEVLCQKQDPDVGIRQRKLGDAWVKRLVEDILSSCKV 651
DB 593 QSDFGKVTMQPFEVLCQKQDPDVGIRQRKLGDAWVKRLVEDILSSCKM 643

RESULT 8
OSU440_MOUSE PRELIMINARY; PRT; 643 AA.
AC OSU440;
DT 07-DEC-2004, integrated into UniProtKB/TrEMBL.
DT 07-DEC-2004, sequence version 1.
DT 07-FEB-2006, entry version 16.
DE Maternal embryonic leucine zipper kinase.
GN Name=Melk;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridea; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Embryo; STRAIN=C57BL/6J;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.P., Casavant T.L., Scheetz T.E.,
RA Raha S.A., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting R.W., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalek U., Smailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Embryo;
RA Director MGC Project;
RA Submitted (OCT-2004) to the EMBL/GenBank/DBJ databases.
CC 1- SIMILARITY: Belongs to the Ser/Thr protein kinase family.
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs license
CC -----
DR EMBL; BC085276; AAH85276.1; -; mRNA.
DR Ensembl; ENSMUSG0000035683; Mus musculus.
DR MGI; MGI:106924; Melk.
DR GO; GO:0005737; Cytoplasm; IDA.
DR GO; GO:0004713; F:protein-tyrosine kinase activity; RCA.
DR InterPro; IPR001772; Kinase_C.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR008271; Ser_thr_pkin_AS.
DR InterPro; IPR002290; Ser_thr_pkinase.
DR InterPro; IPR001245; Tyr_pkinase.
DR Pfam; PF02149; KAI; 1.
```

DR Pfam; PF00069; Pkinase; 1.  
 DR ProDom; PD000001; Prot kinase; 1.  
 DR SMART; SM00220; S TKC; 1.  
 DR PROSITE; PS00107; PROTEIN KINASE ATP; UNKNOWN\_1.  
 DR PROSITE; PS00111; PROTEIN KINASE\_DOM; 1.  
 DR PROSITE; PS00108; PROTEIN KINASE\_ST; UNKNOWN\_1.  
 KW ATP-binding; Kinase; Nucleotide-binding;  
 KW Serine/threonine-protein kinase; Transferase.  
 SQ SEQUENCE 643 AA; 72757 MW; 058893D779BDFE17 CRC64;

Query Match 82.9%; Score 2859; DB 2; Length 643;  
 Best Local Similarity 83.1%; Pred. No. 3.1e-184;  
 Matches 541; Conservative 46; Mismatches 56; Indels 8; Gaps 2;

QY 1 MKDYDELLKYYELHETIGTGGFAKVLACHILGEMVAIKIMDKNTGLSDLPRIKTEIEA 60  
 DB 1 MKDYDELLKYYELHETIGTGGFAKVLACHILGEMVAIKIMDKNKGSLDLPVKTEIDA 60

QY 61 LKVLRHQHCQLHYVLETKANKIPWLEVCYGGELFDYIIISQDRLSEBETRVFRQIVSAV 120  
 DB 61 LKSLRHQHCQLHYVLETKANKIPWLEVCYGGELFDYIIISQDRLSEBETRVFRQILSAV 120

QY 121 AYVHSGYAHARDLKPENLLPDEYHKLIDFGICAKPKGNKDYHLOTCCGSLAYAAPELI 180  
 DB 121 AYVHSGYAHARDLKPENLLPDEYHKLIDFGICAKPKGNKDYHLOTCCGSLAYAAPELI 180

QY 181 QGKSYLGSEADVMSGILLVYLMCGFLPFDDNNVMAlyKKIMRGKYDVPKWLSPSSILLL 240  
 DB 181 QGKSYLGSEADVMSGILLVYLMCGFLPFDDNNVMAlyKKIMRGKYDVPKWLSPSSILLL 240

QY 241 QQMLQVDPKPKRISMKNLLNHPWIMQDYNYPVWOSKNPFIHLDDDCVTELNVHHRNQT 300  
 DB 241 QQMLQVDPKPKRISMKNLLNHPWIMQDYNYPVWOSKNPFIHLDDDCVTELNVHHRNQT 300

QY 301 MEDLISIMQYDHLTATYLLALLAKKPKVRLSLSPSCQASATPFTDIKSNWLSLEDV 360  
 DB 301 MEDLISIMQYDHLTATYLLALLAKKPKVRLSLSPSCQASATPFTDIKSNWLSLEDV 360

QY 361 TASDKNYPVAGLIDYDCEDDLSTGAATPRTSQTKYWTESNGVESKSLTPALCRTPANKL 420  
 DB 361 TASDKNYPVAGLIDYDCEDDLSTGAATPRTSQTKYWTESNGVESKSLTPALCRTPANKL 420

QY 421 KNENYVTPKSAVNEBYFPBPPTPNQNHKREILTPNRYTPSKARNOKLKETPI 480  
 DB 421 KNENYVTPKSAVNEBYFPBPPTPNQNHKREILTPNRYTPSKARNOKLKETPI 480

QY 481 KIPVNSTGTDKMTGTVISPERCRSEVELDINOAHMETPKKCAKVGSLERGLDKVITV 540  
 DB 481 KIPVNSTGTDKMTGTVISPERCRSEVELDINOAHMETPKKCAKVGSLERGLDKVITV 540

QY 541 LTRSKRKGSGARDGPRRLKLYNVTTLVNPDLNLEIMSLPKKHVDVFKQGYTLKCOQ 600  
 DB 541 LTRSKRKGSGARDGPRRLKLYNVTTLVNPDLNLEIMSLPKKHVDVFKQGYTLKCOQ 600

QY 601 QSDFGKVTWQFELEVQOLQKPDVVGIRROBLKGDWYKRLVEDILSSCKV 651  
 DB 601 QSDFGKVTWQFELEVQOLQKPDVVGIRROBLKGDWYKRLVEDILSSCKV 651

QY 651 QSDFGKVTWQFELEVQOLQKPDVVGIRROBLKGDWYKRLVEDILSSCKV 651  
 DB 651 QSDFGKVTWQFELEVQOLQKPDVVGIRROBLKGDWYKRLVEDILSSCKV 651

RESULT 9  
 Q3TU15 MOUSE  
 ID Q3TU15\_MOUSE PRELIMINARY; PRT; 643 AA.  
 AC Q3TU15;  
 DT 11-OCT-2005, integrated into UniProtKB/TrEMBL.  
 DT 11-OCT-2005, sequence version 1.  
 DE 14 days embryo liver cDNA, RIKEN full-length enriched library,  
 DE clone:4432406P16 product:maternal embryonic leucine zipper kinase,  
 DE full insert sequence.  
 GN Name=Wlk;  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;

OC Muroidea; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=C57BL/6J; TISSUE=Liver;  
 RX MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;  
 RA Carninci P., Hayashizaki Y.,  
 RT "High-efficiency full-length cDNA cloning.";  
 RL Methods Enzymol. 303:19-44(1999).  
 RN [2]  
 RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=C57BL/6J; TISSUE=Liver;  
 RX PubMed=16141073; DOI=10.1126/science.1112014;  
 RA Carninci P., Kasukawa T., Katayama S., Gough J., Frith M.C., Maeda N.,  
 RA Oyama R., Ravasi T., Lenhard B., Wells C., Kodzius R., Shimokawa K.,  
 RA Bajic V.B., Brenner S.E., Batalov S., Forrest A.R., Zavolan M.,  
 RA Davis M.J., Wilm M., Balgopal D., Allen J.E.,  
 RA Ambesi-Impombato A., Apweiler R., Aturaliya R.N., Bailey T.L.,  
 RA Bansal M., Baxter L., Beisel K.W., Bersano T., Bono H., Chalk A.M.,  
 RA Chiu K.P., Choudhary V., Christoffels A., Clutterbuck D.R.,  
 RA Crowe M.L., Dalla E., Dalrymple B.P., de Bono B., Della Gatta G.,  
 RA di Bernardo D., Down T., Engstrom P., Fagioli M., Faulkner G.,  
 RA Fletcher C.P., Fukushima T., Furuno M., Futaki S., Gariboldi M.,  
 RA Georgii-Heminger P., Gingeras T.R., Gojobori T., Green R.E.,  
 RA Gustincich S., Harbers M., Hayashi Y., Hensch T.K., Hirokawa N.,  
 RA Hill D., Hummel L., Iacono M., Ikeo K., Iwama A., Ishikawa T.,  
 RA Jakt M., Kanapin A., Katoh M., Kawasawa Y., Kelso J., Kitamura H.,  
 RA Kitano H., Kollias G., Krishnan S.P., Kruger A., Kummerfeld S.K.,  
 RA Kurochkin I.V., Lareau L.F., Lazarevic D., Lipovich L., Liu J.,  
 RA Liuni S., McWilliam S., Madan Babu M., Madera M., Marchionni L.,  
 RA Matsuda M., Matsuzawa S., Miki H., Mignone F., Miyake S., Morris K.,  
 RA Mottagui-Tabar S., Mulder N., Nakano N., Nakaguchi H., Ng P.,  
 RA Nilsson R., Nishiguchi S., Nishikawa S., Nori F., Ohara O.,  
 RA Okazaki Y., Orlandi V., Pang K.C., Pavan W.J., Pavoni G., Pesole G.,  
 RA Petrovsky N., Piazza S., Reed J., Reid J.F., Ring B.Z., Ringwald M.,  
 RA Rost B., Ruan Y., Salzberg S.L., Sandelin A., Schneider C., Sheng Y.,  
 RA Schonbach C., Sekiguchi K., Semple C.A., Seno S., Sessa L.,  
 RA Shibata Y., Shimada H., Shimada K., Silva D., Sinclair B.,  
 RA Sperling S., Stupka E., Sugtara K., Sultana R., Takenaka Y., Taki K.,  
 RA Tammoja K., Tan S.L., Tang S., Taylor M.S., Tegner J., Teichmann S.A.,  
 RA Ueda H.R., van Nimwegen E., Verardo R., Wei C.L., Yagi K.,  
 RA Yamanishi H., Zabarovsky E., Zhu S., Zimmer A., Hide W., Bult C.,  
 RA Grimmond S.M., Tesdale R.D., Liu E.T., Brusic V., Quackenbush J.,  
 RA Wahlestedt C., Mattick J.S., Hume D.A., Kai C., Sasaki D., Tomaru Y.,  
 RA Fukuda S., Kanamori-Katayama M., Suzuki M., Aoki J., Arakawa T.,  
 RA Iida J., Imamura K., Itoh M., Kato T., Kawaji H., Kawagashira N.,  
 RA Kawahina T., Kojima M., Kondo S., Konno H., Nakano K., Ninomiya N.,  
 RA Nishio T., Okada M., Plesky C., Shibata K., Shiraki T., Suzuki S.,  
 RA Tagami M., Waki K., Watanabe A., Okamura-Oho Y., Suzuki H., Kawai J.,  
 RA Hayashizaki Y.;  
 RT "The transcriptional landscape of the mammalian genome.";  
 RL Science 309:1559-1563(2005).  
 RN [3]  
 RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=C57BL/6J; TISSUE=Liver;  
 RX PubMed=16141073; DOI=10.1126/science.1112009;  
 RG RIKEN Genome Exploration Research Group, and Genome Science Group  
 (Genome Network Core Team) and the FANTOM Consortium;  
 RT "Antisense Transcription in the Mammalian Transcriptome.";  
 RL Science 309:1564-1566(2005).  
 RN [4]  
 RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=C57BL/6J; TISSUE=Liver;  
 RX MEDLINE=22354683; PubMed=12466851; DOI=10.1038/nature01266;  
 RA Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,  
 RA Nikaido I., Otsu N., Saito R., Suzuki H., Yamanaka I., Kiyosawa H.,  
 RA Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojobori T.,  
 RA Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,  
 RA Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,  
 RA Blake J.A., Bradt D., Bruscia V., Chothia C., Corbani L.E., Cousins S.,  
 RA Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,  
 RA Gaasterland T., Gariboldi M., Gliss C., Godzik A., Gough J.,  
 RA Grimmond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,

RA Kanai A., Kawai J. H., Kawasawa Y., Kedzierski R. M., King B. L.,  
 RA Nagagaya A., Kurochkin I. V., Lee Y., Lenhard B., Lyons P. A.,  
 RA Maglott D. R., Maltais L., Marchionni L., McKenzie L., Miki H.,  
 RA Nagashima T., Numata K., Okido T., Pavan W. J., Pertea G., Pesole G.,  
 RA Petrovsky N., Pillai R., Pontius J. U., Qi D., Ramachandran S.,  
 RA Ravasi T., Reed J. C., Reed D. J., Reid J., Ring B. Z., Ringwald M.,  
 RA Sadelin A., Schneider C., Sempke C. A., Setou M., Shimada K.,  
 RA Sultana R., Takenaka Y., Taylor M. S., Teasdale R. D., Tomita M.,  
 RA Verardo R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C.,  
 RA Wilming L. G., Wynshaw-Boris A., Yanagisawa M., Yang L., Yang L.,  
 RA Yuan Z., Zavalan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,  
 RA Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,  
 RA Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,  
 RA Hara A., Hashizume W., Iotani K., Ishii Y., Itoh M., Kagawa I.,  
 RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,  
 RA Yasunishi A., Yoshino M., Waterston R., Lander E. S., Rogers J.,  
 RA Birney E., Hayashizaki Y.;  
 RT "Analysis of the mouse transcriptome based on functional annotation of  
 RT 60,770 full-length cDNAs.";  
 RL Nature 420:563-573 (2002).  
 RN [5]  
 RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=C57BL/6J; TISSUE=liver;  
 RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;  
 RA Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Iehii Y.,  
 RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,  
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,  
 RA Saito T., Okazaki Y., Gojibori T., Bono H., Kasukawa T., Saito R.,  
 RA Kadota K., Matsuda H. A., Ashburner M., Batalov S., Casavant T.,  
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Koichiwa H.,  
 RA Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,  
 RA Schriml L. M., Staubli R., Suzuki R., Tomita M., Wagner L., Washio T.,  
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,  
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M. F.,  
 RA Brownstein M. J., Bult C., Fletcher C., Fujita M., Gariboldi M.,  
 RA Gustincich S., Hill D., Hofmann M., Hume D. A., Kamiya M., Lee N. H.,  
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombardi P.,  
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,  
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,  
 RA Suzuki H., Toyooka K., Wang K. H., Weitz C., Whittaker C., Wilming L.,  
 RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawai J. H., Kohseki S.,  
 RA Hayashizaki Y.;  
 RT "Functional annotation of a full-length mouse cDNA collection.";  
 RL Nature 409:685-690 (2001).  
 RN [6]  
 RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=C57BL/6J; TISSUE=liver;  
 RX MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;  
 RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,  
 RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;  
 RT "Normalization and subtraction of cap-trapper-selected cDNAs to  
 RT prepare full-length cDNA libraries for rapid discovery of new genes.";  
 RL Genome Res. 10:1617-1630 (2000).  
 RN [7]  
 RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=C57BL/6J; TISSUE=liver;  
 RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;  
 RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,  
 RA Konno H., Akiyama J., Nishi K., Kutsunai T., Tashiro H., Itoh M.,  
 RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,  
 RA Yamamoto R., Matsumoto H., Sakauchi S., Ikegami T., Kashiwagi K.,  
 RA Fujiwaka S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,  
 RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuura S., Kawai J.,  
 RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;  
 RT "RIKEN integrated sequence analysis (RISA) system-384-format  
 RT sequencing pipeline with 384 multicapillary sequencer.";  
 RL Genome Res. 10:1757-1771 (2000).  
 RN [8]  
 RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=C57BL/6J; TISSUE=liver;  
 RA Arakawa T., Carninci P., Fukuda S., Hashizume W., Hayashida K.,  
 RA Hori F., Iida J., Imanura K., Iotani K., Itoh M., Kanagawa S.,  
 RA Kawai J., Kojima M., Konno H., Murata M., Nakamura N., Ninomiya N.,

RA Nishiyori H., Nomura K., Ohno M., Sakazume N., Sano H., Sasaki D.,  
 RA Shibata K., Shiraki T., Tagami M., Tagami Y., Waki K., Watahiki A.,  
 RA Muramatsu M., Hayashizaki Y.;  
 RL Submitted (APR-2004) to the EMBL/GenBank/DBJ databases.  
 CC -!- SIMILARITY: Belongs to the Ser/Thr protein kinase family.  
 CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>  
 CC Distributed under the Creative Commons Attribution-NoDerivs License  
 CC -----  
 CC EMBL: AK161029; BAE36156.1; -; mRNA.  
 DR MGI: MGI:106924; Melk.  
 DR GO: GO:0005737; Cytoplasm; IDA.  
 DR GO: GO:0004713; F:protein-tyrosine kinase activity; RCA.  
 DR InterPro: IPR001172; Kinase C.  
 DR InterPro: IPR000719; Prot kinase.  
 DR InterPro: IPR008271; Ser\_Thr\_kinase.  
 DR InterPro: IPR002290; Ser\_Thr\_kinase.  
 DR Pfam: PF02149; KAI; 1.  
 DR Pfam: PF00069; Pkinase; 1.  
 DR ProDom: PD000001; Prot kinase; 1.  
 DR SMART: SM00220; S\_TKC\_1.  
 DR PROSITE: PS00107; PROTEIN\_KINASE\_ATP; 1.  
 DR PROSITE: PS00111; PROTEIN\_KINASE\_DOM; 1.  
 DR PROSITE: PS00108; PROTEIN\_KINASE\_ST; UNKNOWN\_1.  
 KW ATP-binding; Kinase; Nucleotide-binding;  
 KW Serine/threonine-protein kinase; Transferase.  
 SQ SEQUENCE 643 AA; 72787 MW; 1092C14709A64966 CRC64;  
 Query Match 82.7%; Score 2852; DB 2; Length 643;  
 Best Local Similarity 82.9%; Pred. No. 9.1e-184;  
 Matches 540; Conservative 46; Mismatches 57; Indels 8; Gaps 2;  
 QY 1 MKDYDELLKYELHETIGTGGFAKVKLACHITGEMVAIKIMDKNTLGLDLPRIKTEIRA 60  
 DB 1 MKDYDELLKYELHETIGTGGFAKVKLACHITGEMVAIKIMDKNTLGLDLPRIKTEIRA 60  
 QY 61 LKNLRHQHICOLYHVLETKANKIFMWLEVCYCGELFDYIISQDRLSEETRVVFRQIVSAV 120  
 DB 61 LKSLRHQHICOLYHVLETKANKIFMWLEVCYCGELFDYIISQDRLSEETRVVFRQIVSAV 120  
 QY 121 AVHSGVYAHROLKPENLLFDYHKLKLDIFGLCAKPKGNKYHLQTCGSLAYAAPELI 180  
 DB 121 AVHSGVYAHROLKPENLLFDENHKLKLDIFGLCAKPKGNKYHLQTCGSLAYAAPELI 180  
 QY 181 QGKSYLGSEADYVSMGILLYVLMCGFLPDDNNVALYKKIMRGKYDVDPKWLSPSSILL 240  
 DB 181 QGKSYLGSEADYVSMGILLYVLMCGFLPDDNNVALYKKIMRGKYDVDPKWLSPSSILL 240  
 QY 241 QQMLQVDPKKRISMKNLNHPIMQDYNYPVEWQSKNPFILHDDDCVTELSVHHRNNRQT 300  
 DB 241 QQMLQVDPKKRISMRNLNHPVMQDYSCPVEMQSKTPLTHLDEDCVTELSVHHRSSRQT 300  
 QY 301 MEDLISLMQYDHLTATYLLLLAKKARGKPVRLRLSLFSFCGQASATPFTDIKNNWSLEDV 360  
 DB 301 MEDLISLMQYDHLTATYLLLLAKKARGKPVRLRLSLFSFCGQASATPFTDIKNNWSLEDV 360  
 QY 361 TASDKNYVAGLDYDWCEDDLSTGAATPQTSQFTKYWTESNGVESKSLTPALCRTPANKL 420  
 DB 357 STSDNDCVAGLDYDWCEDDLSTGAATPQTSQFTKYWTESNGVESKSLTPALCRTPANKL 412  
 QY 421 KKNENYVTPKSAVQNEEYFMFPEPTPVNKNQHKRIILTTPNRYTTPSKARNOCLKETPI 480  
 DB 413 MDKENVCTPKSSVKNBQVFSEPKIPVSKNQYKREIPASPTFRPTFAKARAQCLREAPV 472  
 QY 481 KIPVNSTGTDKLMTGTVISPERCSRVELDLNOAHMEETPKRGAKVFGSLERGLDKVITV 540  
 DB 473 RTPGNSAGADTLTTGTVISPERCSRVELDLNOAHMEETPKRGAKVFGSLERGLDKVITV 532  
 QY 541 LTRSKKSGSARDGPRRLKLHYNTTTLVNPQDLNLEIMSLPKKHVDVFKQGYTLKCO 600  
 DB 533 LTRNKKKSGSARDGPRRLKLHYNTTTLVNPQDLNLEIMSLPKKHVDVFKQGYTLKCO 592



QY 601 QSDPGKVTWQFEVCOLOKPDVVGIRORLKGDAWYKRLVEDILSSCKV 651  
DB 593 QSDFGKVTWQFEVCOLOKPDVVGIRORLKGDAWYKRLVEDILSSCKM 643

RESULT 10  
Q4R6B9 MACFA  
ID Q4R6B9\_MACFA PRELIMINARY; PRT; 457 AA.  
AC Q4R6B9;  
DT 19-JUL-2005, integrated into UniProtKB/TrEMBL.  
DT 19-JUL-2005, sequence version 1.  
DT 07-FEB-2006, entry version 6.  
DE Testis cDNA, clone: Q4R6B9-18422, similar to human maternal embryonic  
DE leucine zipper kinase (MLK1).  
OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
OC Cercopithecoidea; Cercopithecinae; Macaca.  
OX NCBI\_TaxID=9541;  
RN [1]  
RP NUCLEOTIDE SEQUENCE.  
RX PubMed=1594441; DOI=10.1093/molbev/msi187;  
RA Osada N., Hirata M., Tanuma R., Kusuda J., Hida M., Suzuki Y.,  
RA Sugano S., Gojobori T., Shen C.-K.J., Wu C.-I., Hashimoto K.,  
RT "Substitution Rate and Structural Divergence of 5'UTR Evolution:  
RT Comparative Analysis Between Human and Cynomolgus Monkey CDNA's";  
RL Mol. Biol. Evol. 22:1976-1982(2005).  
RN [2]  
RP NUCLEOTIDE SEQUENCE.  
RG International consortium for macaque cDNA sequencing and analysis;  
RT "DNA sequences of macaque genes expressed in brain or testis and its  
RT evolutionary implications";  
RL Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.  
CC -----  
CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>  
CC Distributed under the Creative Commons Attribution-NoDerivs License  
CC -----  
EMBL; AB169266; BAE01356.1; -; mRNA.  
DR GO; GO:0005524; F:ATP binding; IEA.  
DR GO; GO:000166; F:nucleotide binding; IEA.  
DR GO; GO:004674; F:protein serine/threonine kinase activity; IEA.  
DR GO; GO:0016740; F:transferase activity; IEA.  
DR GO; GO:000468; P:protein amino acid phosphorylation; IEA.  
DR InterPro; IPR001772; Kinase C.  
DR InterPro; IPR000719; Prot\_kinase.  
DR Pfam; PF02149; KAI; 1.  
DR ProDom; PD000001; Prot\_kinase; 1.  
DR PROSITE; PS0011; PROTEIN\_KINASE\_DOM; 1.  
KW ATP-binding; Kinase; Nucleotide-binding; Transferase.  
SQ SEQUENCE 457 AA; 52621 MW; CA48A54A164022F0 CRC64;

Query Match 68.1%; Score 2348; DB 2; Length 457;  
Best Local Similarity 97.2%; Pred. No. 6.1e-150;  
Matches 444; Conservative 5; Mismatches 8; Indels 0; Gaps 0;

QY 195 MGILLYVMCGFLPFDDNNMALKYKIMRGYDVPKWLSPSSILLQQMLQVDPKKRISM 254  
DB 1 MGILLYVMCGFLPFDDNNMALKYKIMRGYDVPKWLSPSSILLQQMLQVDPKKRISM 60

QY 255 KNLLNHPWIMQDYNYPVEWQSKNPFHLLDDCCVTELSVHRRNRQTMEDLISLWQYDHLT 314  
DB 61 KNLLNHPWIMQDYNYPVEWQSKNPFHLLDDCCVTELSVHRRNRQTMEDLISLWQYDHLT 120

QY 315 ATYLLLLAKKRGKPVRLRLSFCQASATPFTDIKSNWSLEDVTSAGKNVAGLIDY 374  
DB 121 ATYLLLLAKKRGKPVRLRLSFCQASATPFTDIKSNWSLEDVTSAGKNVAGLIDY 180

QY 375 DWCEDDLSTGAATPRTSQFTKYWTESNGVESKSLTPALCRTPANKLKNKENVYTPKSAVK 434  
DB 181 DWYEDNLSTGAATPQKSTQFTKYWTESNGVESKSLTPALCRTPANKLKNKENVYTPKSAVK 240

QY 435 NEEYFMFPPEPKTPVKNQHKREILTPNRYATPISKARNQCLKETPIKMPVNSTGTDLMT 494

DB 241 NEEYFMFPPEPKTPVKNQHKREILTPNRYATPISKARNQCLKETPIKMPVNSTGTDLMT 300

QY 495 GVISPERRCRSVELDLNQAHMEETPKRGAKGVFGLSRLGLDKVITVLTSTRKRGKSARDGP 554  
DB 301 GVISPERRCRSVELDLNQAHMEETPKRGAKGVFGLSRLGLDKVITVLTSTRKRGKSARDGP 360

QY 555 RLKHLHYNTTTLNPNPDOLINEIMSLPKKIVDFVQKGYTLKCTQSDFGKVTWQFELE 614  
DB 361 RLKHLHYNTTTLNPNPDOLINEIMSLPKKIVDFVQKGYTLKCTQSDFGKVTWQFELE 420

QY 615 VCQLOKPDVVGIRORLKGDAWYKRLVEDILSSCKV 651  
DB 421 VCQLOKPDVVGIRORLKGDAWYKRLVEDILSSCKV 457

RESULT 11  
Q91821 XENLA  
ID Q91821\_XENLA PRELIMINARY; PRT; 651 AA.  
AC Q91821;  
DT 01-NOV-1996, integrated into UniProtKB/TrEMBL.  
DT 01-OCT-2001, sequence version 2.  
DT 07-FEB-2006, entry version 39.  
DE P69B3 (Hypothetical protein).  
OS Xenopus laevis (African clawed frog).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;  
OC Xenopodinae; Xenopus; Xenopus.  
OX NCBI\_TaxID=8355;  
RN [1]  
RP NUCLEOTIDE SEQUENCE.  
RC TISSUE=Unfertilized eggs;  
RA Roghi C., Le Guellec R., Paris J., Couturier A., Philippe M.;  
RL Submitted (OCT-1992) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP NUCLEOTIDE SEQUENCE.  
RC TISSUE=Unfertilized eggs;  
RA Tassan J.P.;  
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP NUCLEOTIDE SEQUENCE.  
RC TISSUE=Oocytes;  
RX MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;  
RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,  
RA Richardson P.;  
RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus  
RT initiative";  
RL Dev. Dyn. 225:384-391(2002).  
RN [4]  
RP NUCLEOTIDE SEQUENCE.  
RC TISSUE=Oocytes;  
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton D.K., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,  
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,  
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahy J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,  
RA Whiting R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
RA Butterfield Y.S.W., Krzywinski M.I., Skalska U., Smalios D.E.,  
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length human  
RT and mouse cDNA sequences";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
RN [5]  
RP NUCLEOTIDE SEQUENCE.



```

Db 123 AYVHSQGYAHRDLKPENLLIDEHNLKLIIDFGLCAKPKGGLDYRLNTCCGSPAYAAPELI 182
QY 181 QGKSYLGSEADVMSGILLVYLMCGFLPDDDDNMALYKIMRGKYDVPKWLSPSSILL 240
Db 183 QKRAYIGSEADINSMGVLVYLLCGFLPDDDDNMALYRRTIRGKAVPKWLSPSSTLL 242
QY 241 QOMLOVDPKKRIISGNLHNPWIMODYNYVVEWQSKNPPFHLDDCCVTELSVHRRNQRT 300
Db 243 NQLLOVDPKKRIITVKHLLSHPLMGQYSDAVQWOSKYPFLGHLDEDCVTELSVPHNSREN 302
QY 301 MEDLISLWQYDHLTATYLLMLAKKARGKPVRLRLSSPCGQASATPTDICKNN-WSLED 359
Db 303 ISALISKWYDQMSATFLLQSKKPKRGKRIHLAIPS-QTGHASTQASIGSEKAMTYED 361
QY 360 V-TASDKNYVAGIIDVDMCDDISTGAATP-----RTSQTKYWTESNGVESKSLT 409
Db 362 VPDSCVDPAFGSMWF---SDAASLPEESPLEBSFLNTHRTKQHSRHDITQLGDMFEFTLST 418
QY 410 PALCRTPANKLKNKENYVTPKSAVKNEEYFMPPEKTPVNNKQHKREILTPNRYTTPSK 469
Db 419 PVTRKVASKKHANKENV-DAESALRNELPFPALPAKPAKSAFSKQTEKQVQGIP--FQAPTS 475
QY 470 ARNQCLKETPIKIPVNSGTGDKLMTGVISPERCRSVELDLNOAHMEETPKRKGAUVFGS 529
Db 476 KESQFTVTPVKPKPTNTCELTATEVLPEKCHSDVLDLNRGHVDSSQKKKAKLFGS 535
QY 530 LERGLDKVITVLTSSKRGSGARDPRRLKLIHYNVTTTLVNPDPQLLNEIMSLPKKHVD 589
Db 536 LERGLDKVITVLTGPKKRSRPGPKLKAHYNVTTQLLNPQLLNEIISVLSSKKQVEY 595
QY 590 VQKGYTLKCOQSDPGKVTQWQFLEVCOLOKPDVVGIRRLRGDAWYKRLVEDILSSC 649
Db 596 VKKGYTLKCOQSDPGRESKMFLEVCRLSKNGAVGIRRLRGDAWYKRLVEDILSSC 655
QY 650 KV 651
Db 656 QV 657

RESULT 13
Q7ZUN5 BRARE
ID Q7ZUN5 BRARE PRELIMINARY; PRT; 676 AA.
AC Q7ZUN5;
DT 01-JUN-2003, integrated into UniProtKB/TrEMBL.
DT 07-FEB-2006, entry version 19.
DE Similar to maternal embryonic leucine zipper kinase.
GN Name=melk;
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Saito R., Watanabe S.;
RL Submitted (APR-2003) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Belongs to the Ser/Thr protein kinase family.
CC Copyrighted by the Uniprot Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
CC EMBL; AB108827; BAC75706.1; -; mRNA.
CC HSSP; P49137; 1NY3.
CC ZFIN; ZDB-GENE-990603-5; melk.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0001666; F:nucleotide binding; IEA.
DR GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR InterPro; IPR001772; Kinase C.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR008271; Ser_thr_pkin_AS.

```

```

DR InterPro; IPR002290; Ser_thr_pkinase.
DR InterPro; IPR001245; Tyr_pkinase.
DR Pfam; PF02149; KAI; 1.
DR Pfam; PF00069; Pkinase; 1.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00220; S_TKc; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; UNKNOWN_1.
KW ATP-binding; Kinase; Nucleotide-binding;
KW Serine/threonine-protein kinase; Transferase.
SQ SEQUENCE 676 AA; 76954 MW; 9E24E4B4E8DE5905 CRC64;

Query Match 55.7%; Score 1920; DB 2; Length 676;
Best Local Similarity 57.6%; Pred. No. 8.2e-121;
Matches 392; Conservative 84; Mismatches 150; Indels 54; Gaps 14;

QY 6 ELKKYELHETIGTGFAKVKLACHILLTGEMVAIKIMDKNTLGSLDLPRIKTETEAALKNLR 65
Db 8 ELKKHYEVETIGSGGFAKVKLGRHKLTGKVAIKIMEKKDGLDDLPVRKIEIEMKNLS 67
QY 66 HQHICOLYHVLLETANKIFMLBYCPGGELFDYIIISODRLSEEBETRVVFRQIVSAVAVHS 125
Db 68 HQHVCELYHVIETTSKIYVLEVCYGGELFDYIIIAKORLSEEBETRVVFRQIISALAYVHS 127
QY 126 QGYAHRDLKPENLLFDYHKLKLIIDFGLCAKPKGNKYHLQTCGGSALAAAPELIQGKSY 185
Db 128 QGYAHRDLKPENLLIDEDHNLKLIIDFGLCAKPKGGLGFELLTCCGSPAYAAPELIQKAY 187
QY 186 LGSEADVMSGILLVYLMCGFLPDDDDNMALYKIMRGKYDVPKWLSPSSILLLOOMLO 245
Db 188 LGSEADVMSGILLVYLLCGFLPDDDDNMALYKIMRGKYDVPKWLSPSSILLLOOMLO 247
QY 246 VDPKRIKIMKLNHPWIMODYNYVVEWQSKNPPFHLDDCCVTELSVHRRNQRTMEDLI 305
Db 248 VDPKRLTVKHLDPWNRGYSTPVETHSKYPLGHIDEDCITEMAVTFKQSKQRTIQLV 307
QY 306 SLWQYDHLTATYLLMLAKKARGKPVRLR-----LSFSFGCGQASATPTD--IKSNWNSLED 359
Db 308 SEWKYDQITATYLLMLAKKRGPRVRLAECVPDIVC-----SPLQDMQLKKSLAFTD 362
QY 360 VTASDKNYVAGIIDVDMCDDISTGAA-----TPTSQF---TKYWTESNGVESK 406
Db 363 DDGVHPVLGSMVFPDDCYDDENPMTPLTPKNTHNTTPMKLYPETTEKWE-----M 416
QY 407 SLTPAL-----CRTPAK-----LKNKNVYTPKSAVKNEEYFMPPEKTPV--NKNQHKR 455
Db 417 AYSPIVHSRPCRQKPERERTKENENLAVPT---DGDVFPALPAPRTPTYNRKVKSNR 473
QY 456 EILTTPNRYTTPS-----KARNQCLKETPIKIPVNSTGT---DKLMTGVISPERCHSVEL 508
Db 474 TVMTTPNHNKSKSEVNGKAGSATKEGSRREVEQQQGGQGLNMLAFSPERRSR--L 531
QY 509 DLNQAHMEETPKRKGAUVFGSLERGLDKVITVLTSSKRGSGARDGPRRLKLIHYNVTTTL 568
Db 532 DLAGCQVDSGQKRGKGVFGSLERGLDKVITVLTSSKRGSGARDGPRRLKLIHYNVTTTL 590
QY 569 VNPDDLNLNEMSLPKKHVDVQKGYTLKCOQSDPGKVTQWQFLEVCOLOKPDVVGIRR 628
Db 591 TNADQVQLQILSEILPEKNDVFKQGYTLKCHTQSDFGKVTQWQFLEVCOLOKPDVVGIRR 650
QY 629 QRLKGDWYKRLVEDILSS 648
Db 651 QRLKGDWYKRLVEDILSS 670

RESULT 14
Q7ZU72 BRARE
ID Q7ZU72 BRARE PRELIMINARY; PRT; 676 AA.
AC Q7ZU72;
DT 01-JUN-2003, integrated into UniProtKB/TrEMBL.
DT 01-JUN-2003, sequence version 1.
DT 07-FEB-2006, entry version 22.

```





**This Page Blank (uspto)**

GenCore version 5.1.9  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: November 21, 2006, 14:04:10 ; Search time 201 Seconds  
(without alignments)  
1480.835 Million cell updates/sec

Title: US-10-656-598-2

Perfect score: 3447

Sequence: 1 MKDYDELLKYYELHETIGT.....KGDWYVKRLVEDILSSCKV 651

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2589679 seqs, 457216429 residues

Total number of hits satisfying chosen parameters: 2589679

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

A\_Geneseq\_8:\*  
1: Geneseqp1980s:\*  
2: Geneseqp1990s:\*  
3: Geneseqp2000s:\*  
4: Geneseqp2001s:\*  
5: Geneseqp2002s:\*  
6: Geneseqp2003as:\*  
7: Geneseqp2003Bs:\*  
8: Geneseqp2004s:\*  
9: Geneseqp2005s:\*  
10: Geneseqp2006s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3447	100.0	651	5 AAB47857	Aab47857 KIAA0175
2	3447	100.0	651	6 ABUS6727	Abus6727 Lung canc
3	3447	100.0	651	6 ABUS7635	Abus7635 Different
4	3447	100.0	651	7 ADBB0563	Adbb0563 Ovarian c
5	3447	100.0	651	7 ADEB38347	Adeb38347 Human pro
6	3447	100.0	651	7 ADN39188	Adn39188 Cancer/an
7	3447	100.0	651	8 ADM72651	Adm72651 Human TAS
8	3447	100.0	651	8 ADM72217	Adm72217 Human TAS
9	3447	100.0	651	8 ADN06036	Adn06036 Antipsori
10	3447	100.0	651	8 ADQ19625	Adq19625 Human sof
11	3447	100.0	651	8 ADU06466	Adu06466 Novel bro
12	3447	100.0	651	9 ADX07297	Adx07297 Cyclin-de
13	3447	100.0	651	9 ADW93570	Adw93570 Human mat
14	3447	100.0	651	9 ADY15428	Ady15428 PRO polytp
15	3447	100.0	651	9 ADY20530	Ady20530 PRO polytp
16	3447	100.0	651	9 AEB57148	Aeb57148 Human mat
17	3447	100.0	651	10 AEF71795	Aef71795 Human gen
18	3447	100.0	656	8 ADX91596	Adx91596 Plant ful
19	3447	100.0	656	8 ADX91597	Adx91597 Plant ful
20	3447	100.0	656	8 ADX91598	Adx91598 Plant ful
21	3260.5	94.6	620	8 ABM83187	Abm83187 Human dia
22	3246.5	94.2	628	8 ABM83195	Abm83195 Human dia
23	3216.5	93.3	612	8 ADU66595	Adu66595 Human kin

#### ALIGNMENTS

RESULT 1

AAB47857

ID AAB47857 standard; protein; 651 AA.

XX AC AAB47857;

XX XX

DT 02-APR-2002 (first entry)

XX XX

DE KIAA0175 protein.

XX XX

KW Antisense; KIAA0175; inhibitor; tumour; P21; P53; chemosensitivity;

KW radiosensitivity; gamma-irradiation; hydroxy urea; cell cycle arrest;

KW sensitization; neoplastic disease; chemotherapy; radiotherapy; cancer.

XX OS Homo sapiens.

XX XX

PN WO200191739-A2.

XX XX

PD 06-DEC-2001.

XX XX

PF 30-MAY-2001; 2001WO-US017644.

XX XX

PR 31-MAY-2000; 2000US-0208435P.

XX XX

PA (CHIR ) CHIRON CORP.

XX XX

PI Wu B, Seeley TW, Williams LT;

XX XX

DR WPI; 2002-122034/16.

XX XX

DR N-PSDB; AAI72250.

XX XX

PT New isolated specific inhibitor useful for decreasing the expression of the inhibitor in a mammalian cell.

XX XX

PS Disclosure; Page 57-59; 60pp; English.

XX XX

CC This protein is encoded by the KIAA0175 cDNA. Antisense oligonucleotides

CC were tested to observe their effect on the kinetics of the KIAA0175

CC transcript. Inhibitors of KIAA0175 may be used for decreasing the

CC expression of KIAA0175 in mammalian cell (preferably tumour cell). They

CC may also be used for decreasing the expression of P21 and P53 in a

CC mammalian cell, for increasing the chemosensitivity and/or

CC radiosensitivity of a mammalian cell by measuring a reduction in gamma-

CC irradiation or hydroxy urea induced P53 or P21 protein levels measuring a

CC reduction in gamma-irradiation or hydroxy urea induced cell cycle arrest

CC and measuring an increase in gamma-irradiation or hydroxy urea induced

CC



CC cell sensitization, and for treating neoplastic disease in a mammal. The  
CC inhibitor provides beneficial improvement of chemo and/or radiotherapy  
CC despite low transfection efficiency (10 - 70%) and/or transient gene  
CC expression. The inhibitor decreases the side effects of the cancer  
XX therapy  
XX  
SQ Sequence 651 AA;

Query Match 100.0%; Score 3447; DB 5; Length 651;  
Best Local Similarity 100.0%; Pred. No. 1.2e-294;  
Matches 651; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKDYDELLKYYELHETIGTGGAFAKVKLACHILTGEMVAIKIMDKNTLGSDLPRIKTEIA 60  
DB 1 MKDYDELLKYYELHETIGTGGAFAKVKLACHILTGEMVAIKIMDKNTLGSDLPRIKTEIA 60

QY 61 LKNLRHQHICQLYHVLETTANKIFMVLEYCPGGELFDYIISQDRLSEETRVVFRQIVSAV 120  
DB 61 LKNLRHQHICQLYHVLETTANKIFMVLEYCPGGELFDYIISQDRLSEETRVVFRQIVSAV 120

QY 121 AYVHSGQYAHARDLKPENLLFDEYHKLKLDLFGICAKPKGNKDYHLOTCCGSLAYAAPELI 180  
DB 121 AYVHSGQYAHARDLKPENLLFDEYHKLKLDLFGICAKPKGNKDYHLOTCCGSLAYAAPELI 180

QY 181 QGKSYLGSEADVWMSGILLVLMCGFLPFDDNNMALYKIMRGKYDVPKWLSPSSILL 240  
DB 181 QGKSYLGSEADVWMSGILLVLMCGFLPFDDNNMALYKIMRGKYDVPKWLSPSSILL 240

QY 241 QQMLQVDPKKRISMKNLLNHPIMQDYNYPVEWQSKNPFHLLDDCVTELSVHHRNROT 300  
DB 241 QQMLQVDPKKRISMKNLLNHPIMQDYNYPVEWQSKNPFHLLDDCVTELSVHHRNROT 300

QY 301 MEDLISLWQYDHLTATYLLLLAKKARGKPVRLSLFSFGQASATPTDIKSNWLSLEV 360  
DB 301 MEDLISLWQYDHLTATYLLLLAKKARGKPVRLSLFSFGQASATPTDIKSNWLSLEV 360

QY 361 TASDKNVAGLIDYDCCEDDLSTGAATPRTSQFTKYWTESNGVESKSLTPALCRTPANKL 420  
DB 361 TASDKNVAGLIDYDCCEDDLSTGAATPRTSQFTKYWTESNGVESKSLTPALCRTPANKL 420

QY 421 KKNENYVTPKSAVKNEEYFMFPBKTPVKNQHKREILTPPNRYTTPSKARNQCLKETPI 480  
DB 421 KKNENYVTPKSAVKNEEYFMFPBKTPVKNQHKREILTPPNRYTTPSKARNQCLKETPI 480

QY 481 KIPVNSTGDKMTGTVISPERCRSEVELDINQAHMEETPKKCAKVGSILRGLDKVITV 540  
DB 481 KIPVNSTGDKMTGTVISPERCRSEVELDINQAHMEETPKKCAKVGSILRGLDKVITV 540

QY 541 LTRSKRKGSAARDGPRRLKLYHNTVTRLVNPDQLLNEIMSILPKKHVDVFQKGYYTLKCQT 600  
DB 541 LTRSKRKGSAARDGPRRLKLYHNTVTRLVNPDQLLNEIMSILPKKHVDVFQKGYYTLKCQT 600

QY 601 QSDFGKVTWQFELEVQOLQKPDVVGIRORLKGDAWYKRLVEDILSSCKV 651  
DB 601 QSDFGKVTWQFELEVQOLQKPDVVGIRORLKGDAWYKRLVEDILSSCKV 651

RESULT 2  
ABU56727  
ID ABU56727 standard; protein; 651 AA.  
AC ABU56727;  
XX  
XX  
DT 02-APR-2003 (first entry)  
DE Lung cancer-associated polypeptide #320.  
XX  
XX Lung cancer-associated polypeptide; cytostatic; emphysema;  
XX antiinflammatory; antiasthmatic; non-small cell lung cancer; atelectasis;  
XX small cell lung cancer; benign lesion; precancerous lesion; bronchitis;  
XX chronic obstructive pulmonary disease; hypersensitivity pneumonitis;  
XX interstitial pulmonary fibrosis; fibrosis; asthma; bronchiectasis.  
XX

OS Unidentified.  
XX  
PN WO200286443-A2.  
XX  
PD 31-OCT-2002.  
XX  
XX 18-APR-2002; 2002WO-US012476.  
XX  
XX 18-APR-2001; 2001US-0284770P.  
PR 18-MAY-2001; 2001US-0290492P.  
PR 09-NOV-2001; 2001US-0332449P.  
PR 13-NOV-2001; 2001US-0350666P.  
PR 29-NOV-2001; 2001US-0334370P.  
PR 12-APR-2002; 2002US-0372246P.  
XX  
PA (EOSB-) EOS BIOTECHNOLOGY INC.  
XX  
PI Aziz N, Murray R;  
XX  
DR WPI; 2003-093161/08.  
DR N-PSDB; ABX76456.  
XX  
PT Detecting a lung cancer-associated transcript in a cell from a patient  
PT for treating lung cancer, by contacting a biological sample from the  
PT patient with a polynucleotide that exhibits increased or decreased  
PT expression in lung cancer.  
XX  
PS Claim 27; Page 436; 453pp; English.  
XX  
XX The invention relates to a method for detecting a lung cancer-associated  
CC transcript in a cell from a patient, comprising contacting a biological  
CC sample from the patient with a polynucleotide that selectively hybridises  
CC to a sequence that is at least 80 % identical to a gene that exhibits  
CC increased or decreased expression in lung cancer samples. Lung cancer-  
CC associated polynucleotides and polypeptides are used for identifying a  
CC compound that modulates a lung cancer-associated polypeptide, for  
CC cancer in a patient and for treating a mammal having lung cancer by  
CC administering a modulatory compound identified. The methods are useful  
CC for treating lung cancer, such as small cell lung cancer, non-small cell  
CC lung cancer or other benign or precancerous lesions, e.g. atelectasis,  
CC emphysema, bronchitis, chronic obstructive pulmonary disease, fibrosis,  
CC hypersensitivity pneumonitis, interstitial pulmonary fibrosis, asthma and  
CC bronchiectasis. The genes, polynucleotides and polypeptides are useful  
CC for diagnostic purposes and as targets for screening for therapeutic  
CC compounds that modulate lung cancer, such as antibodies. Sequences  
CC ABU56408-ABU56745 represent lung cancer-associated polypeptides of the  
CC invention  
XX  
SQ Sequence 651 AA;

Query Match 100.0%; Score 3447; DB 6; Length 651;  
Best Local Similarity 100.0%; Pred. No. 1.2e-294;  
Matches 651; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKDYDELLKYYELHETIGTGGAFAKVKLACHILTGEMVAIKIMDKNTLGSDLPRIKTEIA 60  
DB 1 MKDYDELLKYYELHETIGTGGAFAKVKLACHILTGEMVAIKIMDKNTLGSDLPRIKTEIA 60

QY 61 LKNLRHQHICQLYHVLETTANKIFMVLEYCPGGELFDYIISQDRLSEETRVVFRQIVSAV 120  
DB 61 LKNLRHQHICQLYHVLETTANKIFMVLEYCPGGELFDYIISQDRLSEETRVVFRQIVSAV 120

QY 121 AYVHSGQYAHARDLKPENLLFDEYHKLKLDLFGICAKPKGNKDYHLOTCCGSLAYAAPELI 180  
DB 121 AYVHSGQYAHARDLKPENLLFDEYHKLKLDLFGICAKPKGNKDYHLOTCCGSLAYAAPELI 180

QY 181 QGKSYLGSEADVWMSGILLVLMCGFLPFDDNNMALYKIMRGKYDVPKWLSPSSILL 240  
DB 181 QGKSYLGSEADVWMSGILLVLMCGFLPFDDNNMALYKIMRGKYDVPKWLSPSSILL 240

QY 241 QQMLQVDPKKRISMKNLLNHPIMQDYNYPVEWQSKNPFHLLDDCVTELSVHHRNROT 300  
DB 241 QQMLQVDPKKRISMKNLLNHPIMQDYNYPVEWQSKNPFHLLDDCVTELSVHHRNROT 300



PN WO2002102235-A2.  
XX 27-DEC-2002.  
XX 18-JUN-2002; 2002WO-US019297.  
XX 18-JUN-2001; 2001US-0299234P.  
PR 27-AUG-2001; 2001US-0315287P.  
PR 05-SEP-2001; 2001US-0317544P.  
PR 13-NOV-2001; 2001US-0350666P.  
PR 12-APR-2002; 2002US-0372246P.  
XX (EOSB-) EOS BIOTECHNOLOGY INC.  
XX Mack DH, Gish KC;  
XX WPI; 2003-167431/16.  
DR N-PSDB; AD880562.  
XX  
XX Detecting an ovarian cancer-associated transcript in a cell from a  
PT patient, comprises contacting a biological sample from the patient with a  
PT polynucleotide that hybridizes to an ovarian cancer gene.  
XX  
XX Claim 13; Page 319; 332pp; English.  
XX  
XX The invention relates to a method of detecting an ovarian cancer-  
CC associated transcript in a cell from a patient, by contacting a  
CC biological sample from the patient with a polynucleotide that selectively  
CC hybridizes to a sequence at least 80% identical to any of one of 80  
CC nucleic acid sequences given in the specification. The method is useful  
CC in diagnosing ovarian cancer and in identifying and using agents and/or  
CC targets that inhibit ovarian cancer. The nucleic acid molecule,  
CC polypeptide and the antibody may also be used in detecting ovarian  
CC cancers, monitoring and early detection of relapse following treatment,  
CC monitoring response to therapy, selecting patients for post-operative  
CC chemotherapy or radiation therapy, in selecting mode of therapy,  
CC determining tumour prognosis, early detection of pre-cancerous lesions,  
CC and as vaccines. This sequence corresponds to one of the proteins used  
CC for the detection method of the invention.  
XX  
XX Sequence 651 AA;  
SQ  
Query Match 100.0%; Score 3447; DB 7; Length 651;  
Best Local Similarity 100.0%; Pred. No. 1.2e-294;  
Matches 651; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
1 MKDYDELLKYYELHETIGTGGFAKVKLACHILTGEMVAIKMDKNTLGSDDLPRIKTEISA 60  
1 MKDYDELLKYYELHETIGTGGFAKVKLACHILTGEMVAIKMDKNTLGSDDLPRIKTEISA 60  
61 LKNLRHQHICQLYHVLETANKIFWVLEYCPGGELEFDYIISQDLSEETRVVPRQIVSAV 120  
61 LKNLRHQHICQLYHVLETANKIFWVLEYCPGGELEFDYIISQDLSEETRVVPRQIVSAV 120  
121 AYVHSQGYAHRDLKPNELLFDEYHKLKIDFGLCAKPKGNKYOHLQTCGSLAYAAPELI 180  
121 AYVHSQGYAHRDLKPNELLFDEYHKLKIDFGLCAKPKGNKYOHLQTCGSLAYAAPELI 180  
181 QGKSYLGSEADVMSGILLYVLMCGFLPFDDNNMALYKKIMRGKYDVPKWSLPSILL 240  
181 QGKSYLGSEADVMSGILLYVLMCGFLPFDDNNMALYKKIMRGKYDVPKWSLPSILL 240  
241 QQMLQVDPKRIKSNKLNHPWIMQDYNYPVEWQSKNPFTHLDDCCVTELSVHHRNROT 300  
241 QQMLQVDPKRIKSNKLNHPWIMQDYNYPVEWQSKNPFTHLDDCCVTELSVHHRNROT 300  
301 MEDLISLWQYDHLTATYLLLLAKKARGKPVRLSLSPSCQASATPFTDIKSNWNSLEDV 360  
301 MEDLISLWQYDHLTATYLLLLAKKARGKPVRLSLSPSCQASATPFTDIKSNWNSLEDV 360  
361 TASDKNYPVAGLIIDYDWCEDDLSTGAATPRTSQTKYWTESNGVESKSLTPALCRTPANKL 420  
361 TASDKNYPVAGLIIDYDWCEDDLSTGAATPRTSQTKYWTESNGVESKSLTPALCRTPANKL 420

QY 421 KKNENVYTPKSAVKNEEYFMPEPKTPVKNQKHKEILTTNNRYTPPSKARNQCLKETPI 480  
DB 421 KKNENVYTPKSAVKNEEYFMPEPKTPVKNQKHKEILTTNNRYTPPSKARNQCLKETPI 480  
QY 481 KIPVNSTGTDKLMTGVIISPERRCRSVELDLNQAHMEETPKRGAKVFGSLERGLDKVITV 540  
DB 481 KIPVNSTGTDKLMTGVIISPERRCRSVELDLNQAHMEETPKRGAKVFGSLERGLDKVITV 540  
QY 541 LTRSKRKGSGARDGPRRLKLNHYNTTTRLVNPDQLNEIWSILPKKHVDVFOVKGYTLKCO 600  
DB 541 LTRSKRKGSGARDGPRRLKLNHYNTTTRLVNPDQLNEIWSILPKKHVDVFOVKGYTLKCO 600  
QY 601 QSDFGKVTMQFELEVQQLQKPDVVGIRQRQLKGDWVYKRLVEDILSSCKV 651  
DB 601 QSDFGKVTMQFELEVQQLQKPDVVGIRQRQLKGDWVYKRLVEDILSSCKV 651  
RESULT 5  
ADE38347  
ID ADE38347 standard; protein; 651 AA.  
XX ADE38347;  
XX 29-JAN-2004 (first entry)  
XX Human protein 2089 amino acid sequence.  
XX  
XX tumorigenic disorder; angiogenic disorder; aberrant gene expression;  
XX aberrant protein activity; cytostatic; antithyroid; antidiabetic;  
KW ophthalmological; cancer; breast cancer; colon cancer; lung cancer;  
KW prostatic cancer; Grave's disease; diabetic retinopathy; protein 2089.  
XX Homo sapiens.  
OS  
XX WO2003065006-A2.  
XX 07-AUG-2003.  
XX 30-JAN-2003; 2003WO-US002588.  
XX 31-JAN-2002; 2002US-0353600P.  
PR 15-MAR-2002; 2002US-0364517P.  
PR 09-APR-2002; 2002US-0371075P.  
PR 10-APR-2002; 2002US-0371507P.  
PR 16-APR-2002; 2002US-0372984P.  
PR 19-APR-2002; 2002US-0374194P.  
PR 24-MAY-2002; 2002US-0382995P.  
PR 31-MAY-2002; 2002US-0385023P.  
PR 14-JUN-2002; 2002US-0388533P.  
PR 17-JUN-2002; 2002US-0389395P.  
PR 25-JUN-2002; 2002US-0391324P.  
PR 15-JUL-2002; 2002US-0395944P.  
PR 22-JUL-2002; 2002US-0397726P.  
PR 13-AUG-2002; 2002US-0403046P.  
PR 22-AUG-2002; 2002US-0405155P.  
PR 27-AUG-2002; 2002US-0406361P.  
PR 25-OCT-2002; 2002US-0421195P.  
PR 12-NOV-2002; 2002US-0425456P.  
PR 19-NOV-2002; 2002US-0427626P.  
PR 10-DEC-2002; 2002US-0432122P.  
XX (MILL-) MILLENNIUM PHARM INC.  
PA  
XX Hunter JJ, Macbeth KJ, Tsai F, Lesoon A, Lightcap ES;  
PI Williamson MW, Rudolph-Owen LA;  
XX WPI; 2003-646176/61.  
DR N-PSDB; ADE38346.  
XX  
XX Treating subject having tumorigenic disorder or angiogenic disorder  
PT caused by aberrant polypeptide e.g., N-formylpeptide receptor or nucleic  
PT acid, by administering a modulator.

XX	PS	Disclosure; SEQ ID NO 8; 454pp; English.
XX	CC	This invention relates to a novel method of treating a human subject
CC	CC	having a tumorigenic disorder or angiogenic disorder, caused by aberrant
CC	CC	gene expression or activity of an isolated protein, by administering a
CC	CC	modulator. The modulator may have cytostatic, antithyroid, antidiabetic
CC	CC	or ophthalmological activity. The method is useful for treating a subject
CC	CC	having a tumorigenic or angiogenic disorder, in particular for treating
CC	CC	cancer (for example breast cancer, colon cancer, lung cancer or prostatic
CC	CC	cancer) and, for example, Grave's disease and diabetic retinopathy. The
CC	CC	present sequence is the amino acid sequence of the novel isolated human
CC	CC	protein 2089 of the invention.
XX	XX	Sequence 651 AA;
XX	XX	Query Match 100.0%; Score 3447; DB 7; Length 651;
XX	XX	Best Local Similarity 100.0%; Pred. No. 1.2e-294;
XX	XX	Matches 651; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	1	MKDYDELLKYYELHETIGTGGAFAKVKLACHILTGEMVAIKIMDKNTLGSGLPRIKTEIFA 60
DB	1	MKDYDELLKYYELHETIGTGGAFAKVKLACHILTGEMVAIKIMDKNTLGSGLPRIKTEIFA 60
QY	61	LKNLRHQHICOLYHVLETANKIPWVLEYCPGGELFDYIISQRLSEETRVVFRQIVSAV 120
DB	61	LKNLRHQHICOLYHVLETANKIPWVLEYCPGGELFDYIISQRLSEETRVVFRQIVSAV 120
QY	121	AVVHSGYAHARDLKPNLFLDFEYHKLKIDFGLCAPKGNKDVHLQTCGSLAYAAPELI 180
DB	121	AVVHSGYAHARDLKPNLFLDFEYHKLKIDFGLCAPKGNKDVHLQTCGSLAYAAPELI 180
QY	181	QKSYLGSSEADVMSGILLVLMCGFLPDDDDNMALYKIMRGKYDVPKMLSPSSILL 240
DB	181	QKSYLGSSEADVMSGILLVLMCGFLPDDDDNMALYKIMRGKYDVPKMLSPSSILL 240
QY	241	QOMLVDPKRSKMKLLNHPWIMQDYNYPVEWQSKNPIHLDDDCVTELSVHHRNROT 300
DB	241	QOMLVDPKRSKMKLLNHPWIMQDYNYPVEWQSKNPIHLDDDCVTELSVHHRNROT 300
QY	301	MEDLISLWQYDHLTATYLLALLAKARGKPVRLRLSSFCGQASATPFTDIKNNWSLEDV 360
DB	301	MEDLISLWQYDHLTATYLLALLAKARGKPVRLRLSSFCGQASATPFTDIKNNWSLEDV 360
QY	361	TASDKNYVAGLDYDWCEDDLSTGAATPRTSQTKYWTESNGVESKSLTPALCRTPANKL 420
DB	361	TASDKNYVAGLDYDWCEDDLSTGAATPRTSQTKYWTESNGVESKSLTPALCRTPANKL 420
QY	421	KNKENYTPKSAVKNBEEYFMPPEKTPVKNQKHREILTTPNRYTTPSKARNQCLKETPI 480
DB	421	KNKENYTPKSAVKNBEEYFMPPEKTPVKNQKHREILTTPNRYTTPSKARNQCLKETPI 480
QY	481	KIPVNSTGDKLMTGTVISERRCRSVELDLNOAHMEETPKRGAKVFGSLERGLDKVITV 540
DB	481	KIPVNSTGDKLMTGTVISERRCRSVELDLNOAHMEETPKRGAKVFGSLERGLDKVITV 540
QY	541	LTRSRKSGARDGPRRLKHLHYNTVTRLVNPDQLNEIMSILPKKHVDVQKGYTLKCO 600
DB	541	LTRSRKSGARDGPRRLKHLHYNTVTRLVNPDQLNEIMSILPKKHVDVQKGYTLKCO 600
QY	601	QSDFGKVTMQFELEVQOLQKDPVVGIRRLRGKGDAMVYKRLVEDILSSCKV 651
DB	601	QSDFGKVTMQFELEVQOLQKDPVVGIRRLRGKGDAMVYKRLVEDILSSCKV 651
XX	XX	RESULT 6
XX	XX	ADN39188
XX	XX	ID ADN39188 standard; protein; 651 AA.
XX	XX	AC ADN39188;
XX	XX	DT 17-JUN-2004 (first entry)
XX	XX	XX

DE	XX	Cancer/angiogenesis/fibrosis-related polypeptide, SEQ ID NO:506.
XX	XX	Human; differential expression; cancer; angiogenic disorder;
KW	KW	fibrotic disorder; psoriasis; ischaemia; heart disease; atherosclerosis;
KW	KW	inflammatory disease; autoimmune disease;
KW	KW	retinal neovascularisation syndrome; scarring; uterine fibroid;
KW	KW	detection; diagnosis; prognosis; drug screening; drug targeting;
KW	KW	wound healing; contraception; cytostatic; cardiant; immunomodulatory;
XX	XX	vulnary; gene therapy; vaccine.
OS	OS	Homo sapiens.
XX	XX	WO2003042661-A2.
PN	PD	22-MAY-2003.
XX	XX	13-NOV-2002; 2002WO-US036810.
PF	PF	13-NOV-2001; 2001US-0350666P.
XX	XX	21-NOV-2001; 2001US-0332464P.
PR	PR	29-NOV-2001; 2001US-0334393P.
PR	PR	03-DEC-2001; 2001US-0335394P.
PR	PR	14-DEC-2001; 2001US-0340376P.
PR	PR	08-JAN-2002; 2002US-0347211P.
PR	PR	10-JAN-2002; 2002US-0347349P.
PR	PR	08-FEB-2002; 2002US-0355250P.
PR	PR	13-FEB-2002; 2002US-0356714P.
PR	PR	20-FEB-2002; 2002US-0359077P.
PR	PR	29-MAR-2002; 2002US-036809P.
PR	PR	04-APR-2002; 2002US-0370110P.
PR	PR	12-APR-2002; 2002US-0372246P.
PR	PR	05-JUN-2002; 2002US-0386614P.
PR	PR	16-JUL-2002; 2002US-0396839P.
PR	PR	22-JUL-2002; 2002US-039775P.
PR	PR	22-JUL-2002; 2002US-0397845P.
PR	PR	09-SEP-2002; 2002US-0409450P.
XX	XX	(EOSB-) EOS BIOTECHNOLOGY INC.
PA	XX	Afar D, Aziz N, Ginsburg WM, Gish KC, Glynn R, Hevezi PA;
PI	PI	Mack DH, Murray R, Watson SR, Wilson KE, Zlotnik A;
XX	XX	WPI; 2003-468649/44.
DR	DR	N-PSDB; ADN39187.
XX	XX	Determining the presence or absence of a pathological cell in a patient,
PT	PT	useful for diagnosing, prognosing or treating cancer, comprises detecting
PT	PT	a nucleic acid in a biological sample.
XX	XX	Claim 12; SEQ ID NO 506; 1385pp; English.
PS	PS	The invention relates to nucleic acids and proteins (ADN38683-ADN40064)
XX	XX	whose expression is upregulated or downregulated in specific cancers or
CC	CC	other diseases such as angiogenic or fibrotic disorders, and to methods
CC	CC	of determining the presence or absence of a pathological cell in a
CC	CC	patient by detecting a nucleic acid at least 80% identical to those of
CC	CC	the invention or by detecting a polypeptide of the invention. The
CC	CC	invention also relates to expression vectors and host cells comprising a
CC	CC	nucleic acid of the invention; antibodies which specifically bind a
CC	CC	polypeptide of the invention; use of such antibodies for drug targeting;
CC	CC	methods of screening for modulators of activity or expression of the
CC	CC	polypeptides and nucleic acids. The nucleic acids, polypeptides,
CC	CC	antibodies and methods are useful for diagnosing, prognosing and treating
CC	CC	cancer and other conditions such as psoriasis, ischaemia, heart disease,
CC	CC	atherosclerosis, inflammatory diseases, autoimmune diseases, retinal
CC	CC	neovascularisation syndromes, scarring and uterine fibroids. They may
CC	CC	also be useful in wound healing and in contraception. The present
CC	CC	sequence represents a polypeptide of the invention.
XX	XX	Sequence 651 AA;
XX	XX	Query Match 100.0%; Score 3447; DB 7; Length 651;
XX	XX	Best Local Similarity 100.0%; Pred. No. 1.2e-294;

Matches 651; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY	1 MKDYDELLKYYELHETIGTGGPAKVKLACHILTGEMVAIKMDKNTLGSDDLPRIKTEIEA 60
Db	1 MKDYDELLKYYELHETIGTGGPAKVKLACHILTGEMVAIKMDKNTLGSDDLPRIKTEIEA 60
QY	61 LKNLRHQHICQLYHVLGTANKIFMWLEPCPGGELFDYIISQDRLSEETRVVFRQIVSAV 120
Db	61 LKNLRHQHICQLYHVLGTANKIFMWLEPCPGGELFDYIISQDRLSEETRVVFRQIVSAV 120
QY	121 AYVHSQGYAHRDLKPNELLFDEYHKULIDFGLCAKPGKNQYHLQTCGSLAYAAPELI 180
Db	121 AYVHSQGYAHRDLKPNELLFDEYHKULIDFGLCAKPGKNQYHLQTCGSLAYAAPELI 180
QY	181 QGKSYLGSEADVWSMGILLVYLMCGFLPDDNNMALYKKIMRGKYDVPKWLSPSSILL 240
Db	181 QGKSYLGSEADVWSMGILLVYLMCGFLPDDNNMALYKKIMRGKYDVPKWLSPSSILL 240
QY	241 QQMLQVDPKKRISMKNLLNHPIMQDYNYPVEWQSKNPFILHDDDCVTLSVHHRNNRQT 300
Db	241 QQMLQVDPKKRISMKNLLNHPIMQDYNYPVEWQSKNPFILHDDDCVTLSVHHRNNRQT 300
QY	301 MEDLISLMQYDHLTATYLLLLAKKARGKPVRLRLSSFCGQASATPFTDIKSNWNSLEDV 360
Db	301 MEDLISLMQYDHLTATYLLLLAKKARGKPVRLRLSSFCGQASATPFTDIKSNWNSLEDV 360
QY	361 TASDKNYYAGLIDYDWCEDDLSTGAATPRTSOFTKYWTESNGVESKSLTPALCRTPANKL 420
Db	361 TASDKNYYAGLIDYDWCEDDLSTGAATPRTSOFTKYWTESNGVESKSLTPALCRTPANKL 420
QY	421 KKNENYVTPKSAVKNEEYFMFPPEKTPVNKNQHKREILTPNRYTTPSKARNOCLKETPI 480
Db	421 KKNENYVTPKSAVKNEEYFMFPPEKTPVNKNQHKREILTPNRYTTPSKARNOCLKETPI 480
QY	481 KIPVNSTGTDKLMGTGVISPERRCRSVELDLNQAHMEETPKRGKAVFGSLERGLDKVITV 540
Db	481 KIPVNSTGTDKLMGTGVISPERRCRSVELDLNQAHMEETPKRGKAVFGSLERGLDKVITV 540
QY	541 LTRSKRKGSGARDGPRRLKHLHYNVTTTTLVNPOLLNEIMSLPKKHVDVFQGYTLKCCOT 600
Db	541 LTRSKRKGSGARDGPRRLKHLHYNVTTTTLVNPOLLNEIMSLPKKHVDVFQGYTLKCCOT 600
QY	601 QSDFGKVTMQFLEVCQLQKPDVVGIRRRQRLKGDWVYKRLVEDILSSCKV 651
Db	601 QSDFGKVTMQFLEVCQLQKPDVVGIRRRQRLKGDWVYKRLVEDILSSCKV 651
RESULT 7	
ADM72651	
ID	ADM72651 standard; protein; 651 AA.
XX	
AC	ADM72651;
XX	
DT	17-JUN-2004 (first entry)
XX	
DE	Human TASK110 polypeptide (clone DNA255289).
XX	
KW	TASK; tumour-associated kinase; cytosstatic; tumour;
KW	cell proliferative disorder; cancer; transgenic;
KW	chromosome identification; tissue typing; human; TASK110; enzyme.
XX	
OS	Homo sapiens.
XX	
PN	WO2004024063-A2.
XX	
PD	25-MAR-2004.
XX	
PF	05-SEP-2003; 2003WO-US027886.
XX	
PR	11-SEP-2002; 2002US-0410166P.
XX	
PA	(GETH ) GENENTECH INC.

PI	Davis DP, Desauvage FJ, Wood WI, Zhang Z;
XX	
DR	WPI: 2004-282984/26.
DR	N-PSDB; ADM72650.
XX	
PT	New tumor-associated kinase nucleic acids and polypeptides, useful as
PT	hybridization probes for isolating full length TASK DNA, for generating
PT	transgenic animals, in chromosome identification, or for tissue typing.
XX	
PS	Claim 11; Fig 2; 140pp; English.
XX	
CC	The invention relates to tumour-associated kinase (TASK) polypeptides
CC	(II) and encoding polynucleotides. An antibody, oligopeptide (sRNA) or
CC	organic molecule that binds to (II) is useful for treating a mammal
CC	having a tumour comprising cells expressing (II). Antagonists of TASK are
CC	useful for treating or preventing a cell proliferative disorder (e.g.
CC	cancer) associated with increased expression or activity of (II). The
CC	TASK polynucleotides and polypeptides may be used as hybridization probes
CC	for isolating full length TASK DNA, for generating transgenic animals, in
CC	chromosome identification, or for tissue typing. The present sequence
CC	represents a human TASK110 polypeptide.
XX	
SQ	Sequence 651 AA;
Query Match 100.0%; Score 3447; DB 8; Length 651;	
Best Local Similarity 100.0%; Pred. No. 1.2e-294;	
Matches 651; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY	1 MKDYDELLKYYELHETIGTGGPAKVKLACHILTGEMVAIKMDKNTLGSDDLPRIKTEIEA 60
Db	1 MKDYDELLKYYELHETIGTGGPAKVKLACHILTGEMVAIKMDKNTLGSDDLPRIKTEIEA 60
QY	61 LKNLRHQHICQLYHVLGTANKIFMWLEPCPGGELFDYIISQDRLSEETRVVFRQIVSAV 120
Db	61 LKNLRHQHICQLYHVLGTANKIFMWLEPCPGGELFDYIISQDRLSEETRVVFRQIVSAV 120
QY	121 AYVHSQGYAHRDLKPNELLFDEYHKULIDFGLCAKPGKNQYHLQTCGSLAYAAPELI 180
Db	121 AYVHSQGYAHRDLKPNELLFDEYHKULIDFGLCAKPGKNQYHLQTCGSLAYAAPELI 180
QY	181 QGKSYLGSEADVWSMGILLVYLMCGFLPDDNNMALYKKIMRGKYDVPKWLSPSSILL 240
Db	181 QGKSYLGSEADVWSMGILLVYLMCGFLPDDNNMALYKKIMRGKYDVPKWLSPSSILL 240
QY	241 QQMLQVDPKKRISMKNLLNHPIMQDYNYPVEWQSKNPFILHDDDCVTLSVHHRNNRQT 300
Db	241 QQMLQVDPKKRISMKNLLNHPIMQDYNYPVEWQSKNPFILHDDDCVTLSVHHRNNRQT 300
QY	301 MEDLISLMQYDHLTATYLLLLAKKARGKPVRLRLSSFCGQASATPFTDIKSNWNSLEDV 360
Db	301 MEDLISLMQYDHLTATYLLLLAKKARGKPVRLRLSSFCGQASATPFTDIKSNWNSLEDV 360
QY	361 TASDKNYYAGLIDYDWCEDDLSTGAATPRTSOFTKYWTESNGVESKSLTPALCRTPANKL 420
Db	361 TASDKNYYAGLIDYDWCEDDLSTGAATPRTSOFTKYWTESNGVESKSLTPALCRTPANKL 420
QY	421 KKNENYVTPKSAVKNEEYFMFPPEKTPVNKNQHKREILTPNRYTTPSKARNOCLKETPI 480
Db	421 KKNENYVTPKSAVKNEEYFMFPPEKTPVNKNQHKREILTPNRYTTPSKARNOCLKETPI 480
QY	481 KIPVNSTGTDKLMGTGVISPERRCRSVELDLNQAHMEETPKRGKAVFGSLERGLDKVITV 540
Db	481 KIPVNSTGTDKLMGTGVISPERRCRSVELDLNQAHMEETPKRGKAVFGSLERGLDKVITV 540
QY	541 LTRSKRKGSGARDGPRRLKHLHYNVTTTTLVNPOLLNEIMSLPKKHVDVFQGYTLKCCOT 600
Db	541 LTRSKRKGSGARDGPRRLKHLHYNVTTTTLVNPOLLNEIMSLPKKHVDVFQGYTLKCCOT 600
QY	601 QSDFGKVTMQFLEVCQLQKPDVVGIRRRQRLKGDWVYKRLVEDILSSCKV 651
Db	601 QSDFGKVTMQFLEVCQLQKPDVVGIRRRQRLKGDWVYKRLVEDILSSCKV 651

RESULT 8  
 ADM72217  
 ID ADM72217 standard; protein; 651 AA.  
 XX AC  
 ADM72217;  
 XX DT  
 17-JUN-2004 (first entry)  
 XX DT  
 Human TASK110 polypeptide.  
 DE  
 XX  
 TASK; tumour-associated kinase; cytostatic; tumour antigen;  
 KW cell proliferative disorder; cancer; transgenic; human.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO2004024064-A2.  
 XX  
 PD 25-MAR-2004.  
 XX  
 PF 05-SEP-2003; 2003WO-US027894.  
 XX  
 PR 11-SEP-2002; 2002US-0410166P.  
 XX  
 PA (GETH ) GENENTECH INC.  
 XX  
 PI Desauvage FJ, Wood WI, Zhang Z;  
 XX  
 DR WPI; 2004-282985/26.  
 XX  
 DR N-PSDB; ADM72216.  
 XX  
 XX  
 PT New tumor-associated kinase nucleic acids and polypeptides, useful as  
 PT hybridization probes for isolating full length TASK DNA, for generating  
 PT transgenic animals, in chromosome identification, or for tissue typing.  
 XX  
 PS Claim 12; SEQ ID NO 22; 163pp; English.  
 XX  
 CC The invention relates to new isolated tumour-associated kinase (TASK)  
 CC nucleic acid molecules and encoded polypeptides. Cytostatic. The  
 CC antibody, oligopeptide or organic molecule that binds to the TASK  
 CC polypeptide are useful for treating a mammal having a tumour comprising  
 CC cells expressing the polypeptide. Antagonists of TASK are useful for  
 CC treating or preventing a cell proliferative disorder (e.g. cancer)  
 CC associated with increased expression or activity of TASK polypeptide. The  
 CC TASK polynucleotides and polypeptides may be used as hybridization probes  
 CC for isolating full length TASK DNA, for generating transgenic animals, in  
 CC chromosome identification, or for tissue typing. The present sequence  
 CC represents a human TASK polypeptide.  
 XX  
 XX Sequence 651 AA;  
 SQ

Query Match 100.0%; Score 3447; DB 8; Length 651;  
 Best Local Similarity 100.0%; Pred. No. 1.2e-294;  
 Matches 651; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKDYDELLKYELHETIGTGGAFAKVKLACHILTGEMVAIKMDKNTLGSGLPRIKTEIA 60  
 DB 1 MKDYDELLKYELHETIGTGGAFAKVKLACHILTGEMVAIKMDKNTLGSGLPRIKTEIA 60  
 QY 61 LKNLRHQHICQLVHLETANKIFMWLEYCPGGELFDYIISQDLSEETRVVFRQIVSAV 120  
 DB 61 LKNLRHQHICQLVHLETANKIFMWLEYCPGGELFDYIISQDLSEETRVVFRQIVSAV 120  
 QY 121 AYVHSGYAHRLDKPENLAFDEYHKLKLDIFGLCAKPKGNKDYHLQTCGSLAYAAPELI 180  
 DB 121 AYVHSGYAHRLDKPENLAFDEYHKLKLDIFGLCAKPKGNKDYHLQTCGSLAYAAPELI 180  
 QY 181 QKSYLGSSEADVWMSGILLYVLMCGFLPFDDNNVMALYKIMRGKYDVPKWLSPSSILL 240  
 DB 181 QKSYLGSSEADVWMSGILLYVLMCGFLPFDDNNVMALYKIMRGKYDVPKWLSPSSILL 240  
 QY 241 QOMLQVDPKKRISMKULLNHPIMQDYNYPVEWQSKNPFILHDDDCVTELSVHHRNRQT 300  
 DB 241 QOMLQVDPKKRISMKULLNHPIMQDYNYPVEWQSKNPFILHDDDCVTELSVHHRNRQT 300

QY 301 MEDLISLWQYDHLTATYLLILLAKKARGKPVRLRLSSFCQASATPFTDIKSNWSLEDV 360  
 DB 301 MEDLISLWQYDHLTATYLLILLAKKARGKPVRLRLSSFCQASATPFTDIKSNWSLEDV 360  
 QY 361 TASDKNYVAGLIDYDMCEDDLSTGAATPRTSQFTKYWTESNGVESKSLTPALCRTPANKL 420  
 DB 361 TASDKNYVAGLIDYDMCEDDLSTGAATPRTSQFTKYWTESNGVESKSLTPALCRTPANKL 420  
 QY 421 KKNENYVTPKSAVKNEEYFMFPPEKTPVKNQNOHKREILTPNNRYTTPSKARNQCLKETPI 480  
 DB 421 KKNENYVTPKSAVKNEEYFMFPPEKTPVKNQNOHKREILTPNNRYTTPSKARNQCLKETPI 480  
 QY 481 KIPVNSTGTDKLMGTGVISPERCRSVELDLNOAHMEETPKRGAKVFGSLERGLDKVITV 540  
 DB 481 KIPVNSTGTDKLMGTGVISPERCRSVELDLNOAHMEETPKRGAKVFGSLERGLDKVITV 540  
 QY 541 LTRSKKSGSARDGPRELKLHYNVTTRLVNPDOLLNEIMSLPKKHVDVFQKGYTLKCO 600  
 DB 541 LTRSKKSGSARDGPRELKLHYNVTTRLVNPDOLLNEIMSLPKKHVDVFQKGYTLKCO 600  
 QY 601 QSDFGKVTMQFELEVQLOKQPDVVGIRRORLKGDAMVYKRLVEDILSSCKV 651  
 DB 601 QSDFGKVTMQFELEVQLOKQPDVVGIRRORLKGDAMVYKRLVEDILSSCKV 651

## RESULT 9

ADN06036  
 ID ADN06036 standard; protein; 651 AA.  
 XX AC  
 ADN06036;  
 XX DT  
 01-JUL-2004 (first entry)  
 XX DT  
 Antipsoriatic protein sequence #1174.  
 DE  
 XX antipsoriatic; gene therapy; psoriasis; diagnosis.  
 KW Homo sapiens.  
 XX OS  
 WO2004028479-A2.  
 XX PN  
 08-APR-2004.  
 PD  
 25-SEP-2003; 2003WO-US030907.  
 XX PF  
 25-SEP-2002; 2002US-0414006P.  
 XX PR  
 (GETH ) GENENTECH INC.  
 XX PA  
 Bodary S, Clark H, Jackman J, Schoenfeld J, Williams PM, Wood WI;  
 PI Wu TD;  
 XX  
 DR WPI; 2004-305105/28.  
 DR N-PSDB; ADN06035.  
 XX  
 PT New PRO nucleic acid or polypeptide, useful for preparing a  
 PT pharmaceutical composition for diagnosing or treating psoriasis in a  
 PT mammal.  
 XX  
 PS Claim 9; SEQ ID NO 2431; 3069pp; English.  
 XX  
 CC The invention relates to novel polynucleotide and polypeptides for  
 CC treating psoriasis or a sequence having at least 80% identity to the  
 CC above sequences. The nucleic acid is useful for preparing a composition  
 CC for diagnosing or treating psoriasis in a mammal. This sequence  
 CC corresponds to one of the polypeptides of the invention.  
 XX  
 SQ Sequence 651 AA;

Query Match 100.0%; Score 3447; DB 8; Length 651;  
 Best Local Similarity 100.0%; Pred. No. 1.2e-294;  
 Matches 651; Conservative 0; Mismatches 0; Indels 0; Gaps 0;







CC NO:1246 (Genbank EST W28729) is especially preferred). The method of the  
CC invention is utilized in a kit for determining or predicting whether  
CC patient would be susceptible or resistant to treatment by an agent  
CC modulating cdk activity. The invention also describes a method for  
CC utilizing individualized genetic profiles for treating diseases and  
CC disorders based on patient's response and molecular level, specialized  
CC microarrays comprising the biomarkers described, antibodies directed  
CC against the biomarkers and a cell culture model to identify biomarkers.  
CC The cdk modulator is preferably N-5-[(5-(1,1-Dimethylethyl)-2-  
CC oxazoly]methylthio]-2-thiazolyl-4-piperidine carboxamide, 0.5-L-  
CC tartaric acid salt. Note: The sequence data for this patent did not form  
CC part of the printed specification, but was obtained in electronic format  
CC directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences. This  
CC sequence represents a biomarker used in the method of the invention.  
XX  
SQ Sequence 651 AA;

Query Match 100.0%; Score 3447; DB 9; Length 651;  
Best Local Similarity 100.0%; Pred. No. 1.2e-294;  
Matches 651; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MKDYDELLKYYELHETIGTGGFAKVKLACHILTGEMVAIKIMDKNTLGSGLDLPRIKTEIEA 60  
Db 1 MKDYDELLKYYELHETIGTGGFAKVKLACHILTGEMVAIKIMDKNTLGSGLDLPRIKTEIEA 60  
Qy 61 LKNLRHQHICQLYHVLETANKIPMWLEYCPGGELFDYIISQDRLSEETRVVFRQIVSAV 120  
Db 61 LKNLRHQHICQLYHVLETANKIPMWLEYCPGGELFDYIISQDRLSEETRVVFRQIVSAV 120  
Qy 121 AYVHSGQYAHRLDKPENLLFDEYHKLIDFGICAKPKGNKDYHLQTCGSLAYAAPELI 180  
Db 121 AYVHSGQYAHRLDKPENLLFDEYHKLIDFGICAKPKGNKDYHLQTCGSLAYAAPELI 180  
Qy 181 QGKSYLGSEADVWMSGILLVLMCGFLPFDDNNMALYKIMRGKYDVPKWLSPSSILL 240  
Db 181 QGKSYLGSEADVWMSGILLVLMCGFLPFDDNNMALYKIMRGKYDVPKWLSPSSILL 240  
Qy 241 QQMLQVDPKKRISMKNLLNHPMTMODYNPVEMQSKNPETHLDDDCVTELSVHHRNNROT 300  
Db 241 QQMLQVDPKKRISMKNLLNHPMTMODYNPVEMQSKNPETHLDDDCVTELSVHHRNNROT 300  
Qy 301 MEDLISLQWYDHLTATYLLLLAKKARGKPVRLRLSPFCQASATPFTDIKSNWNSLEDV 360  
Db 301 MEDLISLQWYDHLTATYLLLLAKKARGKPVRLRLSPFCQASATPFTDIKSNWNSLEDV 360  
Qy 361 TASDKNVAAGLIDYDWCEDDLSTGAATPRTSQFTKYWTESNGVESKSLTPALCRTPANKL 420  
Db 361 TASDKNVAAGLIDYDWCEDDLSTGAATPRTSQFTKYWTESNGVESKSLTPALCRTPANKL 420  
Qy 421 XNKENYVTPKSAVKNEBEYFMFPKTPVKNQNHKREILATTPNRYTTPSKARNQCLKETPI 480  
Db 421 XNKENYVTPKSAVKNEBEYFMFPKTPVKNQNHKREILATTPNRYTTPSKARNQCLKETPI 480  
Qy 481 KIPVNSTGTDKLMGTGISPERRCRSVELDLNQAHMEETPKRGAKVFGSLERGLDKVITV 540  
Db 481 KIPVNSTGTDKLMGTGISPERRCRSVELDLNQAHMEETPKRGAKVFGSLERGLDKVITV 540  
Qy 541 LTESKRGSGARDGPRRLKLVNVTTLVNPDPQLLNEIMSLPKKHVDVFQKGVTYLLKQOT 600  
Db 541 LTESKRGSGARDGPRRLKLVNVTTLVNPDPQLLNEIMSLPKKHVDVFQKGVTYLLKQOT 600  
Qy 601 QSDFGKVTQMFOFELEVQQLQKPDVVGVRRQRRLKGDWVYKRLVEDILSSCKV 651  
Db 601 QSDFGKVTQMFOFELEVQQLQKPDVVGVRRQRRLKGDWVYKRLVEDILSSCKV 651

RESULT 13  
ADW93570  
ID ADW93570 standard; protein; 651 AA.  
XX  
AC ADW93570;  
XX  
DT 05-MAY-2005 (first entry)

XX  
DE Human maternal embryonic leucine zipper kinase, MELK.  
XX  
KW Maternal embryonic leucine zipper kinase; enzyme; drug screening;  
KW angiogenesis disorder; antiangiogenic; cardiovascular disease; apoptosis;  
KW cancer; cycostatic; neoplasm; diagnosis; gene therapy.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT Domain 11..263  
FT /label = Kinase\_domain  
FT Domain 602..651  
FT /label = Kinase\_associated\_domain\_1  
XX  
PN WO2005016279-A2.  
XX  
PD 24-FEB-2005.  
XX  
PF 12-AUG-2004; 2004WO-US026231.  
XX  
PR 14-AUG-2003; 2003US-0495193P.  
XX  
PA (EXEL-) EXELIXIS INC.  
XX  
PI Kadyk L, Francis GR, Heuer TS, Lickteig K;  
XX  
DR WPI: 2005-173218/18.  
DR N-PSDB; ADW93565, ADW93566, ADW93569.  
XX  
PT Identifying candidate RAC pathway modulating agents useful for diagnosing  
PT or treating e.g. cancer, comprises screening for agents that modulate the  
PT activity of Maternal Embryonic Leucine Zipper Kinase (MELK).  
XX  
PS Example 1; SEQ ID NO 6; 52pp; English.  
XX  
CC The invention relates to identifying a candidate RAC pathway modulating  
CC agent comprising screening for agents that modulate the activity of  
CC Maternal Embryonic Leucine Zipper Kinase (MELK). Also included are  
CC modulating a RAC pathway of a cell or in a mammalian cell and diagnosing  
CC a disease in a patient. In identifying a candidate RAC pathway-modulating  
CC agent, the assay system comprises cultured cells that express the MELK  
CC polypeptide. The cultured cells additionally have defective RAC function.  
CC The methods are useful for diagnosing or treating cancer or for  
CC identifying modulators of RAC pathway, which may be utilized as  
CC therapeutic targets for disorders associated with defective RAC function,  
CC such as cancer, angiogenic disorder and apoptotic disorders. The present  
CC sequence represents human MELK.

SQ Sequence 651 AA;  
Query Match 100.0%; Score 3447; DB 9; Length 651;  
Best Local Similarity 100.0%; Pred. No. 1.2e-294;  
Matches 651; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MKDYDELLKYYELHETIGTGGFAKVKLACHILTGEMVAIKIMDKNTLGSGLDLPRIKTEIEA 60  
Db 1 MKDYDELLKYYELHETIGTGGFAKVKLACHILTGEMVAIKIMDKNTLGSGLDLPRIKTEIEA 60  
Qy 61 LKNLRHQHICQLYHVLETANKIPMWLEYCPGGELFDYIISQDRLSEETRVVFRQIVSAV 120  
Db 61 LKNLRHQHICQLYHVLETANKIPMWLEYCPGGELFDYIISQDRLSEETRVVFRQIVSAV 120  
Qy 121 AYVHSGQYAHRLDKPENLLFDEYHKLIDFGICAKPKGNKDYHLQTCGSLAYAAPELI 180  
Db 121 AYVHSGQYAHRLDKPENLLFDEYHKLIDFGICAKPKGNKDYHLQTCGSLAYAAPELI 180  
Qy 181 QGKSYLGSEADVWMSGILLVLMCGFLPFDDNNMALYKIMRGKYDVPKWLSPSSILL 240  
Db 181 QGKSYLGSEADVWMSGILLVLMCGFLPFDDNNMALYKIMRGKYDVPKWLSPSSILL 240  
Qy 241 QQMLQVDPKKRISMKNLLNHPMTMODYNPVEMQSKNPETHLDDDCVTELSVHHRNNROT 300  
Db 241 QQMLQVDPKKRISMKNLLNHPMTMODYNPVEMQSKNPETHLDDDCVTELSVHHRNNROT 300

Db 241 QQMLQVDPKRI SMKNLLNHPIMQDYNYPVEWQSKNPFILHDDDCVTELSVHHRNRQT 300  
 QY 301 MEDLISLWQYDHLTATYLLALLAKARGKPVRLRLSSFCGQASATPFTDI KSNNWSLEDV 360  
 Db 301 MEDLISLWQYDHLTATYLLALLAKARGKPVRLRLSSFCGQASATPFTDI KSNNWSLEDV 360  
 QY 361 TASDKNYVAGLIDYDCCEDDLSTGAATPRTSQFTKYWTESNGVSKSLTPALCRTPANKL 420  
 Db 361 TASDKNYVAGLIDYDCCEDDLSTGAATPRTSQFTKYWTESNGVSKSLTPALCRTPANKL 420  
 QY 421 KKNENVYTPKSAVKNEEYFMFPKPTPVNKNQHKREILTPNRYTTPSKARNOCLEKPTPI 480  
 Db 421 KKNENVYTPKSAVKNEEYFMFPKPTPVNKNQHKREILTPNRYTTPSKARNOCLEKPTPI 480  
 QY 481 KIPVNSTGDKLMTGVISPERCRSVELDLNQAHEETPKRGAKVFGSLERGLDKVITV 540  
 Db 481 KIPVNSTGDKLMTGVISPERCRSVELDLNQAHEETPKRGAKVFGSLERGLDKVITV 540  
 QY 541 LTRSKRKGSGARDGPRRLKLHYNVTTTTLVNPDDQLLNEIMSLPKKHVDVFKGYTLKCOQ 600  
 Db 541 LTRSKRKGSGARDGPRRLKLHYNVTTTTLVNPDDQLLNEIMSLPKKHVDVFKGYTLKCOQ 600  
 QY 601 QSDFGKVTMQFELEVQOLQKPDVVGIRRORLKGDAWYKRLVEDILSSCKV 651  
 Db 601 QSDFGKVTMQFELEVQOLQKPDVVGIRRORLKGDAWYKRLVEDILSSCKV 651

RESULT 14  
 ADV15428  
 ID ADY15428 standard; protein; 651 AA.  
 XX AC ADY15428;  
 XX DT  
 XX 05-MAY-2005 (first entry)  
 XX DE PRO polypeptide SEQ ID NO 1234.  
 XX KW Antinflammatory; Immune disorder; Dermatological; Immunosuppressive;  
 KW Antirheumatic; Antiarthritic; Osteopathic; Hemostatic; Antianemic;  
 KW Antithyroid; Antidiabetic; Nephrotropic; CNS-Gen.; Hepatotrophic;  
 KW Virucide; Gastrointestinal-Gen.; Antipsoriatic; Antiasthmatic;  
 XX Antiallergic; ds; gene; diagnosis.  
 XX OS Homo sapiens.  
 XX PN WO2005016962-A2.  
 XX PD 24-FEB-2005.  
 XX PF 11-AUG-2004; 2004WO-US026249.  
 XX PR 11-AUG-2003; 2003US-0493546P.  
 XX PA (GETH ) GENENTECH INC.  
 XX PI Abbas A, Clark H, Ouyang W, Williams MP, Wood WI, Wu TD;  
 XX WPI; 2005-182330/19.  
 XX DR  
 XX PT New nucleic acid encoding PRO polypeptide, useful for diagnosing and  
 PT treating an immune related disorder, e.g. systemic lupus erythematosus,  
 PT rheumatoid arthritis, osteoarthritis, thyroiditis, or diabetes mellitus.  
 PS Claim 8; SEQ ID NO 1234; 158pp; English.  
 CC The invention relates to an isolated nucleic acid encoding a PRO  
 CC polypeptide. The polypeptide, agonist or an antagonist, antibody,  
 CC composition, and method are useful for diagnosing and treating an immune  
 CC related disorder, e.g. systemic lupus erythematosus, rheumatoid  
 CC arthritis. The present sequence represents a DNA encoding a PRO  
 CC polypeptide.  
 XX SQ Sequence 651 AA;

Query Match 100.0%; Score 3447; DB 9; Length 651;  
 Best Local Similarity 100.0%; Pred. No. 1.2e-294;  
 Matches 651; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKDYDELLAKYVELHETIGTGGAFAKVLACHILTGEMVAIKIMDKNTLGSGLPRIKTEIEA 60  
 Db 1 MKDYDELLAKYVELHETIGTGGAFAKVLACHILTGEMVAIKIMDKNTLGSGLPRIKTEIEA 60  
 QY 61 LKNLRHQHICQLYHYVLELTANKIFMWLEYCPGGELFDYIISQRLSEETRVVFRQIVSAV 120  
 Db 61 LKNLRHQHICQLYHYVLELTANKIFMWLEYCPGGELFDYIISQRLSEETRVVFRQIVSAV 120  
 QY 121 AYVHSGYAHRLDKPENLLFDYHKLKLDLFGCAKPKGNKDYHLQTCGSLAYAAPELI 180  
 Db 121 AYVHSGYAHRLDKPENLLFDYHKLKLDLFGCAKPKGNKDYHLQTCGSLAYAAPELI 180  
 QY 181 QGKSYLGSEADVSMGILLVLMCGFLPDDDNVVALYKKIMRGKYDVPKWLSPSSILL 240  
 Db 181 QGKSYLGSEADVSMGILLVLMCGFLPDDDNVVALYKKIMRGKYDVPKWLSPSSILL 240  
 QY 241 QQMLQVDPKRI SMKNLLNHPIMQDYNYPVEWQSKNPFILHDDDCVTELSVHHRNRQT 300  
 Db 241 QQMLQVDPKRI SMKNLLNHPIMQDYNYPVEWQSKNPFILHDDDCVTELSVHHRNRQT 300  
 QY 301 MEDLISLWQYDHLTATYLLALLAKARGKPVRLRLSSFCGQASATPFTDI KSNNWSLEDV 360  
 Db 301 MEDLISLWQYDHLTATYLLALLAKARGKPVRLRLSSFCGQASATPFTDI KSNNWSLEDV 360  
 QY 361 TASDKNYVAGLIDYDCCEDDLSTGAATPRTSQFTKYWTESNGVSKSLTPALCRTPANKL 420  
 Db 361 TASDKNYVAGLIDYDCCEDDLSTGAATPRTSQFTKYWTESNGVSKSLTPALCRTPANKL 420  
 QY 421 KKNENVYTPKSAVKNEEYFMFPKPTPVNKNQHKREILTPNRYTTPSKARNOCLEKPTPI 480  
 Db 421 KKNENVYTPKSAVKNEEYFMFPKPTPVNKNQHKREILTPNRYTTPSKARNOCLEKPTPI 480  
 QY 481 KIPVNSTGDKLMTGVISPERCRSVELDLNQAHEETPKRGAKVFGSLERGLDKVITV 540  
 Db 481 KIPVNSTGDKLMTGVISPERCRSVELDLNQAHEETPKRGAKVFGSLERGLDKVITV 540  
 QY 541 LTRSKRKGSGARDGPRRLKLHYNVTTTTLVNPOLLNEIMSLPKKHVDVFKGYTLKCOQ 600  
 Db 541 LTRSKRKGSGARDGPRRLKLHYNVTTTTLVNPOLLNEIMSLPKKHVDVFKGYTLKCOQ 600  
 QY 601 QSDFGKVTMQFELEVQOLQKPDVVGIRRORLKGDAWYKRLVEDILSSCKV 651  
 Db 601 QSDFGKVTMQFELEVQOLQKPDVVGIRRORLKGDAWYKRLVEDILSSCKV 651

RESULT 15  
 ADY20530  
 ID ADY20530 standard; protein; 651 AA.  
 XX AC ADY20530;  
 XX DT  
 XX 05-MAY-2005 (first entry)  
 XX DE PRO polypeptide SEQ ID NO 6336.  
 XX KW Antinflammatory; Immune disorder; Dermatological; Immunosuppressive;  
 KW Antirheumatic; Antiarthritic; Osteopathic; Hemostatic; Antianemic;  
 KW Antithyroid; Antidiabetic; Nephrotropic; CNS-Gen.; Hepatotrophic;  
 KW Virucide; Gastrointestinal-Gen.; Antipsoriatic; Antiasthmatic;  
 XX Antiallergic; diagnosis.  
 XX OS Homo sapiens.  
 XX PN WO2005016962-A2.  
 XX PD 24-FEB-2005.  
 XX PF 11-AUG-2004; 2004WO-US026249.

Job time : 206 secs

```

XX 11-AUG-2003; 2003US-0493546P.
XX (GETH ) GENENTECH INC.
XX Abbas A, Clark H, Ouyang W, Williams MP, Wood WI, Wu TD;
XX WPI; 2005-182330/19.
XX
XX New nucleic acid encoding PRO polypeptide, useful for diagnosing and
XX treating an immune related disorder, e.g. systemic lupus erythematosus,
XX rheumatoid arthritis, osteoarthritis, thyroiditis, or diabetes mellitus.
XX
XX Claim 8; SEQ ID NO 6336; 158pp; English.
XX
XX The invention relates to an isolated nucleic acid encoding a PRO
XX polypeptide. The polypeptide, agonist or an antagonist, antibody,
XX composition, and method are useful for diagnosing and treating an immune
XX related disorder, e.g. systemic lupus erythematosus, rheumatoid
XX arthritis. The present sequence represents a PRO polypeptide.
XX
XX Sequence 651 AA;
XX
Query Match      100.0%; Score 3447; DB 9; Length 651;
Best Local Similarity 100.0%; Pred. No. 1.2e-294;
Matches 651; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MKDYDELLKYYELHETIGTGGAFAKVKLACHILTGEMVAIKIMDKNTLGSGLPRIKTEIEA 60
DB 1 MKDYDELLKYYELHETIGTGGAFAKVKLACHILTGEMVAIKIMDKNTLGSGLPRIKTEIEA 60
QY 61 LKNLRHQHICQLYHVLLETANKIPWLEYCPGGELFDYIISQDRLSEETRVVFRQIVSAV 120
DB 61 LKNLRHQHICQLYHVLLETANKIPWLEYCPGGELFDYIISQDRLSEETRVVFRQIVSAV 120
QY 121 AYVHSQGYAHRDLKPENLLPDEYHKLIDFGICAKPKGNKYHLQTCGSLAYAAPELI 180
DB 121 AYVHSQGYAHRDLKPENLLPDEYHKLIDFGICAKPKGNKYHLQTCGSLAYAAPELI 180
QY 181 QGKSYLGSEADVMSGILLVYLMCGFLPDDDDNMALYKIMRGKYDVPKWLSPSSILL 240
DB 181 QGKSYLGSEADVMSGILLVYLMCGFLPDDDDNMALYKIMRGKYDVPKWLSPSSILL 240
QY 241 QQMLQVDPKKRI SMKNLLNHPIMQDYNYPVEWQSKNPFILHDDDCVTELSVHRRNRQT 300
DB 241 QQMLQVDPKKRI SMKNLLNHPIMQDYNYPVEWQSKNPFILHDDDCVTELSVHRRNRQT 300
QY 301 MEDLISLWQYDHLTATYLLLLAKKARGKPVRLRLSSFCGQASATPPTDIKSNWLSLEV 360
DB 301 MEDLISLWQYDHLTATYLLLLAKKARGKPVRLRLSSFCGQASATPPTDIKSNWLSLEV 360
QY 361 TASDKNVAAGLIDYDWCEDDLSTGAATPRTSQFTKYWTESNGVESKSLTPALCRTPANKL 420
DB 361 TASDKNVAAGLIDYDWCEDDLSTGAATPRTSQFTKYWTESNGVESKSLTPALCRTPANKL 420
QY 421 KKNENVYTPKSAVKNEEYFMFPBKPTPVNKNQHKREILTPNRYTTPSKARNQCLKETPI 480
DB 421 KKNENVYTPKSAVKNEEYFMFPBKPTPVNKNQHKREILTPNRYTTPSKARNQCLKETPI 480
QY 481 KIPVNSTGTDKLMTGVI SPERRCRSVELDLNQAHMEETPKRGAKVFGSLERGLDKVITV 540
DB 481 KIPVNSTGTDKLMTGVI SPERRCRSVELDLNQAHMEETPKRGAKVFGSLERGLDKVITV 540
QY 541 LTFSKRGSGARDGPRRLKLYHNTTTLVNPDPQLLNEIMSLPKKHVDVFQKGYTLKCO 600
DB 541 LTFSKRGSGARDGPRRLKLYHNTTTLVNPDPQLLNEIMSLPKKHVDVFQKGYTLKCO 600
QY 601 QSDFGKVTMQFELVCOLOKQPDVVGIRORLKGDAWYKRLVEDILSSCKV 651
DB 601 QSDFGKVTMQFELVCOLOKQPDVVGIRORLKGDAWYKRLVEDILSSCKV 651

```

GenCore version 5.1.9  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: November 21, 2006, 14:08:31 ; Search time 44 Seconds  
(without alignments)  
1423.571 Million cell updates/sec

Title: US-10-656-598-2  
Perfect score: 3447  
Sequence: 1 MKDYDELLKYYELHETIGT.....KGDWVYKRLVEDILSSCKV 651  
Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues 283416  
Total number of hits satisfying chosen parameters:  
Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries  
Database : PIR 80.\*  
1: pir1.\*  
2: pir2.\*  
3: pir3.\*  
4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2107	61.1	651	2 S52244	p69Eg3 protein - A
2	1156	33.5	726	2 T33998	hypothetical prote
3	789.5	22.9	713	2 S27966	probable serine/th
4	762.5	22.1	745	2 G01025	serine/threonine p
5	712.5	20.7	774	2 I48609	probable serine/th
6	690.5	20.0	1192	2 T18611	probable serine/th
7	690.5	20.0	1246	2 G89287	protein H39E23.1 l
8	648	18.8	513	1 S60303	serine/threonine-s
9	647	18.8	552	1 S51025	[hydroxymethyl]glut
10	644	18.7	552	1 A53621	[hydroxymethyl]glut
11	641.5	18.6	504	2 T10449	probable serine/th
12	640.5	18.6	504	2 T07415	probable serine/th
13	637	18.5	512	1 JCI446	serine/threonine-s
14	637	18.5	513	1 S60304	serine/threonine-s
15	636	18.5	602	2 S72513	POG2 protein - yea
16	635	18.4	633	1 A26030	serine/threonine-s
17	630	18.3	512	2 T52633	serine/threonine-s
18	627	18.2	480	2 A86427	probable serine/th
19	627	18.2	562	2 T29858	hypothetical prote
20	625	18.1	472	2 B90100	SNF-related kinase
21	624.5	18.1	511	1 A56009	serine/threonine-s
22	624	18.1	891	2 T40503	protein kinase kin
23	623	18.1	891	2 A38903	protein kinase 1 -
24	622	18.0	441	2 C84667	probable protein k
25	618.5	17.9	798	2 JCY500	qik protein - chic
26	615.5	17.9	473	1 S59941	serine/threonine-s
27	611	17.7	502	1 A41361	serine/threonine-s
28	610.5	17.7	512	2 T07788	probable serine/th
29	608.5	17.7	887	2 T20941	hypothetical prote

30 608 17.6 510 2 T04145 serine/threonine p  
31 606.5 17.6 440 2 T14736 probable serine/th  
32 606.5 17.6 440 2 T14735 probable serine/th  
33 605.5 17.6 445 2 T50802 serine/threonine p  
34 605.5 17.6 481 2 I49072 protein kinase - m  
35 605 17.6 1398 2 T13741 hypothetical prote  
36 601.5 17.4 576 2 T41587 probable carbon ca  
37 595 17.3 445 2 T09903 serine/threonine-s  
38 593 17.2 435 2 E84707 probable protein k  
39 593 17.2 502 2 T02306 probable protein k  
40 592 17.2 489 2 T04862 probable serine/th  
41 592 17.2 746 2 S62365 SNF1-related prote  
42 588.5 17.1 622 1 S44859 serine/threonine-s  
43 587.5 17.0 442 2 T48203 hypothetical prote  
44 579 16.8 426 2 C71408 probable protein k  
45 578.5 16.8 453 2 G86141 protein T25K16.13

ALIGNMENTS

RESULT 1  
S52244  
p69Eg3 protein - African clawed frog  
C:Species: Xenopus laevis (African clawed frog)  
C:Date: 07-May-1995 #sequence\_revision 03-Aug-1995 #text\_change 05-Oct-2004  
C:Accession: S52244  
R:Roghi, C.; le Guelllec, R.; Paris, J.; Couturier, A.; Philippe, M.  
submitted to the EMBL Data Library, October 1992  
A:Description: Eg3, selected by differential screening encodes a new Xenopus protein kinase  
A:Reference number: S52243  
A:Accession: S52244  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-651 <ROG>  
A:Cross-references: UNIPROT:Q91821; UNIPARC:UPI000017A463; EMBL:Z17205; NID:g609283; PID:Z17205  
C:Keywords: ATP  
F:11-265/Domain: protein kinase homology <KIN>  
F:19-27/Region: protein kinase ATP-binding motif

Query Match	61.1%	Score	2107;	DB	2;	Length	651;
Best Local Similarity	62.0%	Pred. No.	2.6e-85;				
Matches	410;	Conservative	87;	Mismatches	138;	Indels	26;
Gaps	8;						
QY	3	DYDELLKYYELHETIGTGGFAKVKLACHILITGEVVAIKIMDKNTLGSIDLPRKTEIALK	62				
Db	5	DYDELLKYYELHETIGTGGFAKVKLASHLITGEVVAIKIMDKESLGDDLPRVKTEIDAMK	64				
QY	63	NLRHQHICOLYHVLETANKIFMWLEPCGGELFDYIIISQDRLSEBETRVVFRQIVSAVAY	122				
Db	65	NLSHQHVCLXHVIEPTPKKIFMWLEPCGGELFDYIIAKDRLTEEARVFFRQIVSAVAY	124				
QY	123	VHSQGVYARHDLAKPENLLPDEYHKLKLDIFGLCAKPKGNKDYHLOTCCGSLAAVAAPLIGQ	182				
Db	125	IHSQGVYARHDLAKPENLLPDEYHKLKLDIFGLCAKPKGGLDYHLMTCGSPAAAPLIGQ	184				
QY	183	KSYLGSEADVWSMGILLVYLMCGFLPFDNDNVYALYKKIMRGKYDVPKWLSPSSILLLOQ	242				
Db	185	KAYIGSEADVWSMGVLMYALMGVLPFDNDNVYALYKKIMRGKYEPKWLSPGVLLLSQ	244				
QY	243	MLQVDPKGRKISMKNLNHPWIMQDYNYPVEWQSKNPFHLDDDCVTELSVHHRNRQTWE	302				
Db	245	MMQVDPKGRITVKKHLNHPWLMHGYSCPVWQSKNPLYGVIDEDCVTELSVFYRTSTT	304				
QY	303	DLISLWQVDHLLTATYLLALLAKKARGKPVRLRLSSFCGQASATPFTDIKNNWSLEDVTA	362				
Db	305	RLISEWSYDHIITASYLLHLSKSHKAVLKL-HPLAVGQAVTSFKELAPKSLDFEEN	363				
QY	363	SDKNVTAGLIDY---DWCEDDLSTGAATPRT-SQFTKKWTSNGVESKSLTLPALCRTPAN	418				
Db	364	GBIAYVFGSMDSDELSEDFITYSSFEPTPKYVKGRLFNVSVDSPATPVFNAMLGT	423				
QY	419	KLNKENVYTPKSAVK-----NEEYFMPPEPTKPVNKNQHKREILTTNRYTTPSKARNQ	473				

```
Db 424 KMKIKRTVITQWPGMKWMSFCTQHRGLLP-----GGRMKRRVY-SPLLIKTLKQKRKI 476
QY 474 CLKETPIK-----IPVNSTGTDKMTGTVISPBRRCSRVELDLNOAHMETPRKRGAKVFGSL 530
Db 477 SOKRKQKSLVLP-----GEEFANVISPBRRCSRVELDLNOAHIDSQAQKKGAKVFGSL 530
QY 531 ERGLDKVITVLTSTRSKSGARDPRRLKLHYNVTTLRLVNPDLQLLNEIMSILPKKHVDVF 590
Db 531 ERGLDKMITLTSKRGKGYTREGPRKLRAHYNVTTLNINVPDLQLLNQIVRVLPSKKNVDYV 590
QY 591 QKGYTLKCTQSDFGKVTMQFELEVQLQKQSDPVDVGIRRQRLKGDVWYKRLVEDILSSCK 650
Db 591 QKGYTLKCTQSDFGKVTMQFELEVQLQKQSEVVGIRRQRLKGDVWYKRLVEDILSSCK 650
QY 651 V 651
Db 651 V 651

RESULT 2
T33998
Hypothetical protein W03G1.6 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 05-Oct-2004
C:Accession: T33998
R:Pauley, A.; Scheet, P.; Harper, M.
submitted to the EMBL Data Library, February 1999
A:Description: The sequence of C. elegans cosmid W03G1.
A:Reference number: Z21454
A:Accession: T33998
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-726 <PAU>
A:Cross-references: UNIPROT:Q9UAY1; UNIPARC:UPI0000082A03; EMBL:AF125964; PIDN:AAD14754.
A:Experimental source: strain Bristol N2; clone W03G1
C:Genetics:
A:Gene: CESP:W03G1.6
A:Map position: 4
A:Introns: 21/1; 48/3; 87/3; 224/3; 280/3; 347/3; 468/3; 523/1; 584/1; 697/3

Query Match 33.5%; Score 1156; DB 2; Length 726;
Best Local Similarity 37.8%; Pred. No. 9.7e-44;
Matches 281; Conservative 121; Mismatches 230; Indels 112; Gaps 19;

QY 1 MKDYDELLKYELHETIGTGGPAKVKLACHILTGEMVAIKIMDKNTLGSDDLPRIKTEIEA 60
Db 1 MSKYEVLOQGYAVHDLGSGGFGKVRATHLLTNQKVAIKIIDKQLGHDLPVQTEMDA 60
QY 61 LKNLRHOICOLYHVLLETANKIFWVLEYCPGGLFDYIISODRLSEETRVVFRQIVSAV 120
Db 61 LRNLSHQICRLHYHLETEDKFFIVMEYCGGEMFDYIVRKRLERSEARHFRQIVSAI 120
QY 121 AYVHSGYAHRLDKPENLLPDEYHKLKLIDFGLCAK-PKGNKDYH-LQTCGSGLAYAAPE 178
Db 121 AFVHSGYAHRLDKPENLLATEDLHLKLIDFGLCAKTEKGRIDKHNLDTCGSPAYAAPE 180
QY 179 LIQCKSYLGEADVWSMGILLYVLMCGFLPFDDDDNVMALYKIMRGKYDVPKWLSPSSIL 238
Db 181 LIQGLQYKGEADVWSMGILLYTLVGALPFDDNNQIMYKIKQSCFPYEPFELSLSKQ 240
QY 239 LLOQMLQVDPKKRISKMNLLNHPWIMQDYNYPVWQSKNPFILHDDDCVTLSVHHR-NN 297
Db 241 LLRAMLOVPERISVKKLLSHDLNHNKYTOPKWNVTYIDKFNIDRDVARVMSKYYGFES 300
QY 298 ROTMEDLISLWQYDHLTATYLLLLAKKARGKPVRLRLSSFCQQA-----SATPFT 348
Db 301 TDKNWIEKIKWNFDYMTSTYVALLHRKNGWEIILPMVRNSTWTAPNVQNILCSPTHA 360
QY 349 DIKSNWNSLEDVTASDKNYVAGLIDYDWCEDDLSTGAATPRTSQFTKYWTESNGVES--- 405
Db 361 SLENN-----LDK---SGLEDDD--SDPSSISGSSDISARLKKNCVVSDESSSRF 406
```

```
QY 406 -KSLTPALCR-----TPANKLNKKNENVYTPKS--AVKNBEYFMFPBPKTP----- 447
Db 407 VKPMSPAAEKDKKMSVVMNMLTMPQFTGRSLRIPESPMVSRSSDSASLGSAATPSRG 466
QY 448 -----VNQNH-----KREILTTNRYTTPSKARNQCLKETP 479
Db 467 VKDNDKENASTGKNYRMGASTCKSRGPKLITGVGAENVTIKTQYLTLLKISRENYKEGT 526
QY 480 IKIPVNSTGTDKMTGTVISP-----BRRCRSVE-----LDLNOAHME----- 516
Db 527 MGSVYTPNTRTLRGLFSPGNAEHKKQARASDRASIGMPGSPVSGSAHSANNADG 586
QY 517 ETPK-----RKGAUVFSGSLERGLDKVITVLTSTRSKSGARDGPRRL-----KLHYNVTYTR 567
Db 587 RTPRSRIKTNRLPQRFVTSLEKKEKLIITLTPRKMQ--RDSPOVLKDVKNMVNVSMTA 643
QY 568 LVNPDLQLLNEIMSILPKKHVDVFQKGYTLKCTQSDPFGKVTMQFELEVQLQKQSDPVGIR 627
Db 644 SQDPBEVRNLLKKVDFDERMYELNGWKFLATQETVHGMMTV--ELEIVRLQMFQDKVGIR 701
QY 628 RORLKGDAWYKRLVEDILSSCKV 651
Db 702 RKRLKGDAFWYKVKCEKILQMAKI 725

RESULT 3
S27966
probable serine/threonine-specific protein kinase (EC 2.7.1.-) - human
N:Alternate names: protein p78
C:Species: Homo sapiens (man)
C:Date: 17-Apr-1993 #sequence_revision 17-Apr-1993 #text_change 05-Oct-2004
C:Accession: S27966
R:Maheshwari, K.K.; Som, S.; Parsa, I.
submitted to the EMBL Data Library, January 1992
A:Description: Sequence of a cDNA encoding 78KD marker protein lost in chemically induced
A:Reference number: S27966
A:Accession: S27966
A:Molecule type: mRNA
A:Residues: 1-713 cMAH>
A:Cross-references: UNIPROT:P27448; UNIPARC:UPI0000047FOC; EMBL:M80359; NID:g189511; PIDN:
C:Keywords: ATP; phosphotransferase; serine/threonine-specific protein kinase
F:54-307/Domain: protein kinase homology <KIN>
F:62-70/Region: protein kinase ATP-binding motif

Query Match 22.9%; Score 789.5; DB 2; Length 713;
Best Local Similarity 30.6%; Pred. No. 9.6e-28;
Matches 217; Conservative 119; Mismatches 254; Indels 119; Gaps 22;

QY 11 YELHETIGTGSPAKVKLACHILTGEMVAIKIMDKNTLG--SDLPRIKTEIALKNLRHQHI 69
Db 56 YRLKLTIGKGNPAKVKLARHILTGREVAIKIIDKTLNPTSLQKLPREVIMKILNHPNI 115
QY 70 COLYHVLLETANKIFWVLEYCPGGLFDYIISODRLSEETRVVFRQIVSAVAVHSGYVA 129
Db 116 VKLFEVETQKTYLTIMYASGGKVFYLVARGMRKEKARSFRQIVSAVQVCHOKRIV 175
QY 130 HRDLKPENLLPDEYHKLKLIDFGLCAK-PKGNKDYHLOTCCGSLAYAAPELQKSYLGS 188
Db 176 HRDLKAENLLDADNNIKIADFGSNEFTVGK---LDTFCGSPPYAAPELFGKKYDGP 232
QY 189 RADVMSGILLYVLMCGFLPFDDDDNVMALYKIMRGKYDVPKWLSPSSILLQOMLQVDP 248
Db 233 EVDVMSGLVILTVLSGSLPFDGQNLKELRERVLRGKRYIPFVMSDTCENLLKRFVLNLP 292
QY 249 KGRISKMNLLNHPWIMQDYNYPVWQSKNPFTH-----LDDDCVTLSVHHRNNRQTMEDL 304
Db 293 IKRGTLQIMKDRWI-----NAGHEEDLPFPPELIDSDQKRIDIMVGKGYSQEIQBS 348
QY 305 ISLWQYDHLTATYLLLLAKKARGKPVRLRLSSFCQQAATPFTDI-----KSNWNS 356
Db 349 LSKMKIDEITATYLLLRGRASSEVRP-----SDLLNNSTQSPHHKVRQSVSSQKORYS 403
QY 357 -----LEDVTASDKNYVAGLIDYDWCEDDLSTGAATPRTSQFTKYWTESNGVESKSLTP 410
```

```
Db 404 DHAGPQIPSVAYPKRSQSTADSDLKEDGISRSKST-----GSAVGGKGIAP 451
Qy 411 A--LCRTPAKLGK-----NKENVYT-----PKSAVKN--EE 437
Db 452 ASPMLGNASPNKADIPERKKSSTVPSSNTASGTMTRTYVCSERTTDDRHSVIQNGKE 511
Qy 438 YMFPEPKPTPVNKNQKHREILTT--PNRYTTPSKARNQC-----LKETPIKIPVNSTGTDKLM 493
Db 512 NSTIPDQRTFV--ASTHSISAAATPDR-----IRFP-----RGTSASRTFHG 551
Qy 498 SP-ERRC-----RSVELDLNQAHEETPKRKGAVFGSLERGLDKVITVLTSS----- 544
Db 552 QPRERTATYNGPPASPSPLSHETATPUSQTRSRSSTILF-----SKLTSKLTSSRNWSA 604
Qy 545 KRKGSARDG--PRRLKLHLYNVTTTTLVNPDPQLNEIMSILPKKHVDVQK--GYTLKQQTOS 602
Db 605 KQDENKEAKPSRLRFTWSMKTTSSMDPGDMREIRKVLNCDVQERERFLFCVHOD 664
Qy 603 DFGKVTMQFELEVQCOLQKPDVVGIRQRORLKGDAWVYKRLVEDILSSCKV 651
Db 665 GHAENLVQWEMEVCKLPRLSLNGVRPKRISGTSIAPFNIAASKIANELKL 713

RESULT 4
G01025
serine/threonine protein kinase - human
C:Species: Homo sapiens (man)
C>Date: 21-Dec-1996 #sequence_revision 06-Jun-1997 #text_change 05-Oct-2004
C:Accession: G01025
R:Navarro, E.
submitted to the EMBL Data Library, April 1996
A:Reference number: H00564
A:Accession: G01025
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-745 <NAV>
A:Cross-references: UNIPROT:Q15524; UNIPROT:Q96HB3; UNIPARC:UPI000017A42E; EMBL:X97630;
F:18-271/Domain: protein kinase homology <KIN>

Query Match 22.1%; Score 762.5; DB 2; Length 745;
Best Local Similarity 29.6%; Pred. No. 1.5e-26;
Matches 220; Conservative 136; Mismatches 267; Indels 121; Gaps 27;

Qy 11 YELHETIGTGGAFAKVKLACHILITGEMVAIKIMDKNTL--GSDLPRIKTEIALKNLRHOHI 69
Db 20 YRLAKTIGKGNFAKVKLARHILITGKAVKIDKTLQNLSSSLQKLFREVRIMKVLNHPNI 79
Qy 70 COLYHVLTEANKIFWVLEPCPGGELFDYIIISODRLSEBETRVVFRQIVSAVAYVHSQGYA 129
Db 80 VKLFEVETETKTLVLMVEYASGGEVFDYLVAGHRMKEKEARAKFRQIVSAVQYCHQKFTV 139
Qy 130 HRDLKAPENLLPDEYHKLIIDFGLCAK--PKGNKDYHLOTCCGSLAYAAPELLQKSYLGS 188
Db 140 HRDLKAENLLDADNMIKIADFGSNEFTFGNK---LDTFCGSPPYAAPELFGQKKYDGP 196
Qy 189 EADVSMGILLVLMCGFLPFDDDNVMAIYKIMRGKYDVPKWLSPSILLQOVLQVDP 248
Db 197 EVDVWSLGVILTYLVSGSLPFQGNLRELRLVLRGKTRIPFYMSTDCENLLKKFLINLP 256
Qy 249 KGRISMKNLLNHPIMQDYNYPVWQSKNPFTH-----LDDCCVTLSVHRRNNRQTMEDL 304
Db 257 SKRGTLQIMKDRW---NVGHEDDELKPYVEPLPDYKPRTELWMSMGVTREEIQDS 312
Qy 305 ISLWQYDHLTATYILLAKKA--RGKPVRLR-----LSSPSC-----QOASATP-- 346
Db 313 LVQRYNEVMATYLLGVKSSELEGGDTTLKPRPSADLTNSSAQPFPHKVRQSVSANPKQ 372
Qy 347 --FTD-----IKSNWSLEDVTASDKNVVAGLIDYDWCEDDLSTCAATPTSTQPTKYWT 398
Db 373 RRFSDQAGPAIPTSNYSYKKTOSNAENKRP-----BEDRESGRKASSTAKVPA--S 422
Qy 399 ESNGVESKSLTP-----ALCRTPANKLNK-----ENVYTPKSAVKN--EYFMPPEPK----- 445
```

```
Db 423 PUPGLERKKTTPPTSTNSVLSTSTNRSRNSPLLERASLQASIQNGKOSLTPFGSRSTA 482
Qy 446 -----TPVNKNQKHREILTT--PNRYTTPSKARNQC-----LKETPIKIPVNSTGTDKLM 493
Db 483 SASAGVSAARPRQHQKSMGSHVHPNKAISGLPFTESNCEVPRSTAPQRVPVVASPSAHNIS 542
Qy 494 TQVISPERS-----RCRSVELDLNQAHEETPK-----RKGA--KVP 527
Db 543 SSGGAPDRTNFPRGVSSRSRSTFHAGQLRQVR--DQONLPYGVTPASPSPSGHSGQRRGASGSI 601
Qy 528 GS-----LERGL-----DKVITVLT-----RKRKGSARDG--PRRLKLHLYNVTTTTL 568
Db 602 SKFTSKFVRNLINEPESKDRVETLPHVVGSGNDKEKEEFREAPKRSURFTWSMKTTSS 661
Qy 569 VNPDPQLNEIMSILPKKHVDV--VQKGYTLKQQTOSDFQKVTMQFELEVQCOLQKPDVVGIR 627
Db 662 MEPNEMMEIRKVLNCDANSQSELHEKYMLLCMHGTGPHGDFVQWEMEVCKLPRLSLNGVR 721
Qy 628 RQRLKGDWVYKRLVEDILSSCKV 651
Db 722 FKRISGTSWAFKNIAASKIANELKL 745

RESULT 5
I48609
probable serine/threonine-specific protein kinase (EC 2.7.1.-) kem - mouse
C:Species: Mus musculus (house mouse)
C>Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 05-Oct-2004
C:Accession: I48609; S31333
R:Ingalls, J.D.; Lee, M.; Hill, R.E.
Mamm. Genome 4, 401-403, 1993
A:Title: Emk, a protein kinase with homologs in yeast maps to mouse chromosome 19.
A:Reference number: I48609; MUID:93364122; PMID:8358177
A:Accession: I48609
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-774 <ING1>
A:Cross-references: UNIPROT:Q05512; UNIPARC:UPI00000275D5; EMBL:X70764; NID:g57919; PIDN:
submitted to the EMBL Data Library, January 1993
R:Ingalls, J.D.; Lee, M.; Hill, R.E.
A:Description: A novel protein kinase with homologues in yeast maps to mouse chromosome 19.
A:Reference number: S31333
A:Accession: S31333
A:Molecule type: mRNA
A:Residues: 1-698, 'K', 700-702, 'GRGLPTAAKASCTGTCYCAGMAHQATRT', 731-774 <ING2>
A:Cross-references: UNIPARC:UPI000017A424; EMBL:X70764
C:Genetics:
A:Gene: emk
C:Keywords: ATP; phosphotransferase; serine/threonine-specific protein kinase
F:51-304/Domain: protein kinase homology <KIN>
F:59-67/Region: protein kinase ATP-binding motif

Query Match 20.7%; Score 712.5; DB 2; Length 774;
Best Local Similarity 28.6%; Pred. No. 2.4e-24;
Matches 210; Conservative 138; Mismatches 282; Indels 103; Gaps 25;

Qy 11 YELHETIGTGGAFAKVKLACHILITGEMVAIKIMDKNTL--GSDLPRIKTEIALKNLRHOHI 69
Db 53 YRLAKTIGKGNFAKVKLARHILITGKAVKIDKTLQNLSSSLQKLFREVRIMKVLNHPNI 112
Qy 70 COLYHVLTEANKIFWVLEPCPGGELFDYIIISODRLSEBETRVVFRQIVSAVAYVHSQGYA 129
Db 113 VKLFEVETETKTLVLMVEYASGGEVFDYLVAGHRMKEKEARAKFRQIVLHVQYCHQKFTV 172
Qy 130 HRDLKAPENLLPDEYHKLIIDFGLCAK--PKGNKDYHLOTCCGSLAYAAPELLQKSYLGS 188
Db 173 HRDLKAENLLDADNMIKIADFGSNEFTFGNK---LDTFCGSPPYAAPELFGQKKIDGP 229
Qy 189 EADVSMGILLVLMCGFLPFDDDNVMAIYKIMRGKYDVPKWLSPSILLQOVLQVDP 248
Db 230 EVDVWSLGVILTYLVSGSLPFQGNLRELRLVLRGKTRIPFYMSTDCENLLKKFLINLP 289
```



QY 249 KKRISMKNLLNHPWIMQDYNYPVEMQSKNPFIFHL-----DDCVTELSEVHHNNRQTMED 303  
DB 290 SKRGTLLEQIMKDRM-----NVGHEDDELKPYVEPLTTGPRDRVDGVNGLH---TEEIQD 342  
QY 304 LISLMQYDHLTATYLLLLAKKA--RQKVRLL--RLSSFGCGQASATPFTDI--KSNWNLSLE 358  
DB 343 SLVGQRYNEVMATYLLGLYKSSPEGDITTLKPRPSADLTNSAPSCHKVORSVSNPK 402  
QY 359 DVTSADKNVYVAGLIDYDWCEDDLSTGAATPRTSQFTKYWTESN-----GVESKSILT 409  
DB 403 QRRSSDQAVPAIPTSNYSKKTQSNNAENKRPEEETGRKASSTAKVPASPLPGLDKRKT 462  
QY 410 PA-----LCRTFANKLKNK---ENVYTPKSAVKN--BEYFMFPEPK-----TPVNK 450  
DB 463 PAPSTNSVLSTSNRSNRPFLDRASLGQASIQNGKDSLTPGSRASTASAAVSAARP 522  
QY 451 NQHKREILTT--PNRYTTPSKARNQC-----LKETPIKIPVNSTGTDKLMGTGVISPER---- 501  
DB 523 RQHKQMSASVHPNKASGLPPTESNCEVPRPSTAPQRPVPSAHNISSSSGAPDRTNF 582  
QY 502 -----RCSVELDLNQAHWEETPK-----RKGA--KVFGS-----LERG 533  
DB 583 PRGVSSRSTFHAGQLRQVR--DOQLPYGVTPASPSPGSHSGRRGASGSIKFTSKFVRN 641  
QY 534 L-----DKVITVLTR-----SKRKGSGARDG--PRRLKLNHNVVITRLVNPQDLINEM 579  
DB 642 LNEPESKDRVETLRPHVVGSGGTDKKEEFREAKPRSLRFTWSMKTTSMEPNEMWREIR 701  
QY 580 SILPKKHVPF--VQGYTLKQCTQSDGKVTMQEVLVCQLOKQPDVVVGIRRQRLKGDWVY 638  
DB 702 KVLDA NSCQSELHERYMLLCVHGTPGHENFVQWEMEVCKLPRLSLNGVRFKRIISGTSMAF 761  
QY 639 KRLVEDILSCKV 651  
DB 762 KNIASKIANELKL 774  
RESULT 6  
T18611  
N;Contains: probable serine/threonine-specific protein kinase (EC 2.7.1.-), long splice form - Caenorhabditis  
C;Species: Caenorhabditis elegans  
C;Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004  
C;Accession: T18611; T18610; T23144; T23143  
R;McMurray, A.  
submitted to the EMBL Data Library, October 1996  
A;Reference number: Z18997  
A;Accession: T18611  
A;Status: translated from GB/EMBL/DBJ  
A;Molecule type: DNA  
A;Residues: 1-1192 <W11>  
A;Cross-references: UNIPROT:Q9TW45; UNIPARC:UPI000007ECA3; EMBL:Z81027; PIDN:CAB54179.1;  
A;Experimental source: Clone AH10  
A;Accession: T18610  
A;Status: translated from GB/EMBL/DBJ  
A;Molecule type: DNA  
A;Residues: 1-487,536-1192 <W12>  
A;Cross-references: UNIPARC:UPI000017BD4F; EMBL:Z81027; PIDN:CAB54178.1; GSPDB:GN000023;  
A;Experimental source: Clone AH10  
R;McMurray, A.  
submitted to the EMBL Data Library, June 1997  
A;Reference number: Z19696  
A;Accession: T23144  
A;Status: translated from GB/EMBL/DBJ  
A;Molecule type: DNA  
A;Residues: 1-1192 <W13>  
A;Cross-references: UNIPARC:UPI000007ECA3; EMBL:Z96102; PIDN:CAB54263.1; GSPDB:GN000023;  
A;Experimental source: Clone H39E23  
A;Accession: T23143  
A;Status: translated from GB/EMBL/DBJ  
A;Molecule type: DNA  
A;Residues: 1-487,536-1192 <W14>  
A;Cross-references: UNIPARC:UPI000017BD4F; EMBL:Z96102; PIDN:CAB54262.1; GSPDB:GN000023;

A;Experimental source: clone H39E23  
C;Genetics:  
A;Gene: CESP:H39E23.1a; CESP:H39E23.1b  
A;Map position: 5  
A;Introns: 10/3; 33/3; 67/3; 139/2; 183/3; 451/3; 487/3; 535/3; 631/1; 825/2; 914/3; 992/3  
C;Keywords: alternative splicing; ATP; phosphotransferase; serine/threonine-specific prot  
F;1-1192/Product: probable serine/threonine-specific protein kinase, long splice form #st  
F;1-487,536-1192/Product: probable serine/threonine-specific protein kinase, short splice  
Query Match 20.0%; Score 690.5; DB 2; Length 1192;  
Best Local Similarity 31.7%; Pred. No. 3.3e-23;  
Matches 185; Conservative 85; Mismatches 173; Indels 141; Gaps 16;  
QY 3 DYDELLKTYELHETIGTGCFARVKLACHILTGEMVAIKMDKNTLG--SDLPRIKTEIAL 61  
DB 162 DQDVHVGVKYLKLTIGKGNFAVKLAKHVITGHEVAIKIIDKTALNPSSLQKLPREVKIM 221  
QY 62 KMLRHOHICOLYHVHLETKANIPMLVEYCPGGBLFDYIISQDLSEBEETRVFRQIVSA 121  
DB 222 KQLDHPNIVKLYQVMEETQTLVLEAYASGGVFYDLVAHGRMKKEAKARFQIVSAVQ 281  
QY 122 VVHSQGYAHRDLKPNLLFDEYHKLKLDIFGLC--AKPKGNKDYHLQTCGSLAYAAPELI 180  
DB 282 YLHSGNIIHRDLKAENLLDQDMNIIADFGSNTFSLGNK---LDTPCGSPPYAAPELP 338  
QY 181 QGKSYLGSSEADVWSMGILLYVLMCGFLPFDDDDNNVALYKIMRGKYDVPKWLSPSILL 240  
DB 339 SGKKYDGPEDVWSLGVILYTLVSGSLPFDGQNLKELRERVLRGKVRIPFYMSTDCENLL 398  
QY 241 QOMLQVDPKKRISMKNLLNHPWIMQDYNYPVEMQSKNPFIFHLDDDCVTELSEVHHNNRQT 300  
DB 399 KKFLVINPQRRSSLDNIMKDRMW---NVGYDDEDLKPFIEPPKQIDF-----QR 445  
QY 301 MEDLISLWQ-----YDHLTATYLL----- 320  
DB 446 IEKLIQIFQLGPNKAAILESVEKEKFEIDHATYLLLGERKSDMDASEITMAOSLLSHSI 505  
QY 321 -----LAKKARGKPVRLRLSSPSCQASATPPTDIKSNWNLSLEDVTASDKNVAGLI---- 372  
DB 506 NVSSSLGQHPAGVITREHVTVSSASGSSASP-----SRYSSSATATCASITAGSALAS 559  
QY 373 -----DYDWCEDDLSTGAATPRTSQ-----FTKYWTESNGVE-----SKSLT 409  
DB 560 AANAQKHQSSAAPSAGSSSSSSRRSSQNDAAATAAGTVVMSGTRHGQVQMAQRTSROAT 619  
QY 410 PALCRTPANK-----LKNKNVYTPKSAVKNEEYFMPEPEKTPVKNQKHKE 456  
DB 620 ISLLQPPSYKPSNSTTQIAIPLPFRNSTATSSAA-----QPSTGI----- 661  
QY 457 ILTTPNRYTPSKARNQCLKETPIKIPVNSTGTD--KLMTGVIS 498  
DB 662 ---TGTRKIADPKGR-----IPLNSTAVQGHRTATGAVA 692  
RESULT 7  
G89287  
protein H39E23.1 [imported] - Caenorhabditis elegans  
C;Species: Caenorhabditis elegans  
C;Date: 10-May-2001 #sequence\_revision 10-May-2001 #text\_change 10-May-2001  
C;Accession: G89287  
R;anonymous, The C. elegans Sequencing Consortium.  
A;Title: Genome sequence of the nematode C. elegans: a platform for investigating biology  
Science 282, 2012-2018, 1998  
A;Reference number: A75000; MUID:99069613; PMID:9851916  
A;Note: see websites genome.wustl.edu/gsc/C\_elegans/ and www.sanger.ac.uk/Projects/C\_eleg  
A;Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and  
A;Accession: G89287  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-1246 <STO>  
A;Cross-references: UNIPARC:UPI000017A692; GB:chr\_V; PIDN:CAB09532.1; PID:G3878100; GSPDB  
C;Genetics:  
A;Gene: H39E23.1

A:Map position: 5

Query Match 20.0%; Score 690.5; DB 2; Length 1246;  
Best Local Similarity 31.7%; Pred. No. 3.5e-23;  
Matches 185; Conservative 85; Mismatches 173; Indels 141;

```

Qy   3 DYDELLKYELHETITGCGPAKVAKLACHILTGEMVAIKIMDKNTLG--SDIPRIKTEIAL 61
Db   52 DQDVHVGVGYKLLKTIGKNFAKVAKLAHVITGHEVAIKIIDKTALNPSSLQKLFPREVKIM 111
Qy   62 KNLRAHQHCOLYHVLELTANKIWMVLEYCPGGELFDYIIISODRLSEEBTRVVPRQIVSAVA 121
Db  112 QOLDHPNIVKUQYWEETEQTUYLEVASGSVEVDYLVAHGMMKEKGARAKFRQIVSAVQ 171
Qy  122 YVHSQGYAHRDLKPENLLFDEYHKLIKLIIDFLC- AKPQGNKDYLHOTCCGSLAYAAPBLI 180
Db  172 YLHKSNIITHRDLKAENLLLDQDMNIKIADFGSNFTPSLGNK---LDTFCGSPPYAAPELF 228
Qy  181 OGKSYLGSEADVWSMGILLYYVMCGPFLPFDDNDNMALYKKIMRGKYOVPKWLSPLSLLL 240
Db  229 SGKXYDGPEDVMSGLVILYTLVSGSLFFPDQONLKELRERVLRGKYRIPFYMSTDCEPLL 288
Qy  241 QOMLOVPKKIRSMKNLLNHFWIMQDNYPVPMOSKNPFIHLDDCVTELSSHVHNRRQT 300
Db  289 KKFLVIMPQRASSLONIWKORWM---NVGYEDELAPFEPPKQIDF-----QR 335
Qy  301 MEDLISLWQ-----YDHLTATYLLL----- 320
Db  336 IEKLIQIFQLGFNKAAILESVEKEKPEDIHATYLLGERKSDMDASEITMAQSILLSHSI 395
Qy  321 -----LAKARGKPVURLSFSFCQASATPFTDIKNNWSLEDVTASDKNYVAGLI---- 372
Db  396 NVSSSLGQHAGVITREHVITSSSSAGSSASP-----SRYSRSATATGASITAGSALAS 449
Qy  373 -----DYDWCEDDLSTGAATPTSQ-----FTKYWTESNVB-----SKSLT 409
Db  450 AANAQKHQQSSAAPSSGSSSSRRSSQNDAAATAAGTVVMSGTHKGVGWRAQPTSQAT 509
Qy  410 PALCKRTPANK-----LNKENVYTPKSAVNBEYFMFPBPKTPVANKOHKRE 456
Db  510 ISLLQPPSYKPSNSTTTQIAIQIPLFNENSTATSSAA-----QPSTGI----- 551
Qy  457 ILTPPNRYTPPSKARNOCCLKETPIKIPVNSTGTD--KLMTGVIS 498
Db  552 ---TGTRKIADPKGR-----IFLNSFVGOGHRTATCAVA 582

```

## RESULT 8

S60303  
serine/threonine-specific protein kinase (EC 2.7.1.-) BKIN12 (version 1) - barley  
C/Species: Hordeum vulgare (barley)  
C/Date: 20-Jul-1996 #sequence\_revision 27-Feb-1997 #text\_change 05-Oct-2004  
C/Accession: S60303; S24578  
R/Halford, N.G.; Vicente-Carabajosa, J.; Sabelli, P.A.; Shewry, P.R.; Hannappel, U.; Krejci  
Plant J. 2, 791-797, 1992  
A/Title: Molecular analyses of a barley multigene family homologous to the yeast protein  
A/Reference number: S60303; MUID:93258420; PMID:1302632  
A/Accession: S60303  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-513 <HAL>  
A/Cross-references: UNIPROT:Q40029; UNIPARC:UP000000A154A; EMBL:X65506; NID:g18931; PIDN  
C/Genetics:  
A/Introns: 62/1; 126/3; 187/3; 231/3; 293/3; 323/3; 351/3; 398/3; 476/3  
C/Function:  
A/Description: catalyzes the formation of peptidyl-serine-phosphate or peptidyl-threonin  
C/Superfamily: SNF1-related protein kinase; protein kinase homology  
C/Keywords: ATP; magnesium; phosphotransferase; serine/threonine-specific protein kinase  
F/15-272/Domain: protein kinase homology <KIN>  
F/23-31/Region: protein kinase ATP-binding motif  
F/46,65,143,145/Active site: Lys, Glu, Asp, Lys #status predicted  
F/148,152/Binding site: magnesium (Asn, Asp) #status predicted

Query Match 18.8%; Score 648; DB 1; Length 513;  
Best Local Similarity 38.6%; Pred. No. 1.1e-21;

Matches	151; Conservative	68; Mismatches	128; Indels	44; Gaps
Qy	4	YDELLKYVELHETITGTPGPAKVAKLACHILITGEMVAIKIMDKNTLUGSDLPRIKTIEALKN	63	
Db	10	HSEVLKYNLNGKTGLAGTGFQDVKVAEHKLTGQRVAIKILNRKMET-----MEMEKANRE	65	
Qy	64	LR-----HQHICOLHYVLETANKIEMWLEYCPGGBELFDYIISQDRLSSEBETRVVFR	114	
Db	66	IKIMRLFDIIPHIIIRVYQVETPKDIFIVMEYCNGGELLDYIENGRLQGEDEARRIFQ	125	
Qy	115	QIVSAVAVHSQGYAHRDLKPENLLFDYEHKLKLIIDFGLCAKPKGNKDYH-LQTCCGSLA	173	
Db	126	QILAGVEYCHRIWMVHRDLKPENLLDLSKYNVKLADFGL---SNVMRDGHFLTKSCGSLN	182	
Qy	174	YAAPBLTQKSYLGSSEADVMSMGLLIYVLMCGFLPFDDDDNVMALYKKIMRGKYDVPKWL	233	
Db	183	YAAPEIISSKLYAGPEYDVMSGCVILYALLCGSVPFDDDNIPSLFRKIKGGTYTILPSYLS	242	
Qy	234	PSSTILLQOOLQVDPKKRISMNLLNHPW-----IMQDVNYVPVQSKNPFTHLD	283	
Db	243	DSARDLPLKLNLDIPMKRIITHEIRVHPWFKNHLPCVLAVPPPYKAP---KAK-----MID	295	
Qy	284	DDCVTELSVHRRNNRQTMEDLISIQVDHLTATYLLLLAKKAKGPKVRLRLRSLSPSCGOAS	343	
Db	296	EDILRDVVNLGYDKDHCVESLNWRLQNEETVAYVLLL-----DNRFRSTGVLG-AD	346	
Qy	344	ATPPTDIKSNWSL-EDVTASDKNYVAGLID	373	
Db	347	HOHLMDRSFNEFTLSESASPTSNYLPNGID	377	

## RESULT 9

S51025  
 [hydroxymethylglutaryl-CoA reductase (NADPH2)] kinase (EC 2.7.1.109) - human  
 N:Alternate names: AMPK  
 N:Contains: [acetyl-CoA carboxylase] kinase (EC 2.7.1.128); [hydroxymethylglutaryl-CoA reductase] (EC 2.7.1.128)  
 C:Species: Homo sapiens (man)  
 C:Date: 01-Aug-1995 #sequence\_revision 01-Sep-1995 #text\_change 05-Oct-2004  
 C:Accession: S51025; I38503  
 R:Berri, R.K.; Marley, A.E.; See, C.G.; Sopwith, W.F.; Aguan, K.; Carling, D.; Scott, J.; FEBS Lett. 356, 117-121, 1994  
 A:Title: Molecular cloning, expression and chromosomal localisation of human AMP-activated protein kinase  
 A:Reference number: S51025; MUID:95080410; PMID:7988703  
 A:Accession: S51025  
 A:Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-552 <BER>  
 A:Cross-references: UNIPROT:P54646; UNIPARC:UPI00001250A9  
 R:Aguan, K.; Scott, J.; See, C.G.; Sarkar, N.H.  
 Gene 149, 345-350, 1994  
 A:Title: Characterization and chromosomal localization of the human homologue of a rat AMP-activated protein kinase  
 A:Reference number: I38503; MUID:95047501; PMID:7959015  
 A:Accession: I38503  
 A:Status: preliminary; translated from GB/EMBL/DDBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-179, 'T', 181-270, 'G', 272-402, 'RQ', 405-552 <RES>  
 A:Cross-references: UNIPARC:UPI0000161B8A; EMBL:U06454; NID:g758366; PIDN:AAA64745.1; PII:1000000000  
 C:Genetics:  
 A:Gene: GDB:PRKAA2; PRKAA; AMPK; hAMPK  
 A:Cross-references: GDB:451905; OMIM:600497  
 A:Map position: lp31-lp31  
 C:Function:  
 A:Description: catalyzes the formation of peptidyl-serine-phosphate or peptidyl-threonine-phosphate  
 A:Note: in vivo substrates include hydroxymethylglutaryl-CoA reductase (NADPH) and acetyl-CoA  
 C:Superfamily: SNF1-related protein kinase; protein kinase homology  
 C:Keywords: ATP; magnesium; phosphotransferase; serine/threonine-specific protein kinase  
 F:14-268/Domain: protein kinase homology <KIN>  
 F:22-30/Region: protein kinase ATP-binding motif  
 F:45, 64, 139, 141/Active site: Lys, Glu, Asp, Lys #status predicted  
 F:144, 148/Binding site: magnesium (Aen, Asp) #status predicted

Query Match	18.8%; Score 647; DB 1; Length 552;
Best Local Similarity	43.3%; Pred. No. 1.2e-21;
Matches 139; Conservative	51; Mismatches 115; Indels 16; Gaps 6;
QY	10 YYELHETIGTGGFAKVKLACHILTGMVAIKMDKNTLGS--DLPRKTEIEALKNLRHQ 67
DB	15 HYLVDGTLGVGTGKVKIGEHQHTGHKVAVKILNRQKIRSLDVGKIKREIQNLKLFRRP 74
QY	68 HICQLYHVLLETANKIPMWLEFCPGGELFDYIIISQDRLSEETRVPRQIVSAVAVHSHQ 127
DB	75 HIKLQVISTPTDFPMWYVSGGELFDYICKGRVEVEARLFFQQLSAVDYCHRM 134
QY	128 YAHRLDKPENLLFDEYHKLKLDIFGLCAKPKGNKDYHLQTCGSLAYAAPELIQKSYLG 187
DB	135 VVRDLKPEVLLDAHNAKIADFGLSNMMSDGE--FLRTSCGSPNYAAPEVISGRLYAG 192
QY	188 SEADVWSMGLLYVLMCGFLPFDDNNMVALYKIMRGKYDVPKWLSPSSILLLOQMLOYD 247
DB	193 PEVDIWSGCVILYALLCGTLPPDDEHVPTLFKIRGGVFYIPEYLNRSVATLLMHMLQYD 252
QY	248 PKKRISMKNLNHPWIMQD---YNYVPEWQSKNPFTH--LDDDCVTELSVHHR-NNRQT 300
DB	253 PLKRTATKIDIREHWFQDLPVLPF-----EDPSYDANVIDEAVKEVCEKPECTESEV 307
QY	301 MEDLISLWQYDHLTATYLLLL 321
DB	308 MNSLYSGDPQDLAVAYHLII 328
RESULT 10	
A53621	[hydroxymethylglutaryl-CoA reductase (NADPH2)] kinase (EC 2.7.1.109) - rat
N:Contains:	[acetyl-CoA carboxylase] kinase (EC 2.7.1.128); [hydroxymethylglutaryl-CoA x
C:Species:	Rattus norvegicus (Norway rat)
C:Date:	07-Jul-1995 #sequence_revision 07-Jul-1995 #text_change 05-Oct-2004
C:Accession:	A53621; S53729
R:Carling, D.; Aguan, K.; Woods, A.; Verhoeven, A.J.M.; Beri, R.K.; Brennan, C.H.; Sidel	
J. Biol. Chem.	269, 11442-11448, 1994
A:Title:	Mammalian AMP-activated protein kinase is homologous to yeast and plant protein
A:Reference number:	A53621; MUID:94209324; PMID:7908907
A:Accession:	A53621
A:Status:	preliminary
A:Molecule type:	mRNA
A:Residues:	1-552 <CAR>
A:Cross-references:	UNIPROT:Q09137; UNIPARC:UPI00001250AB; GB:229486; NID:9488375; PIDN:
R:Gao, G.; Widmer, J.; Stapleton, D.; Teh, T.; Cox, T.; Kemp, B.E.; Witters, L.A.	
Biochim. Biophys. Acta	1266, 73-82, 1995
A:Title:	Catalytic subunits of the porcine and rat 5'-AMP-activated protein kinase are
A:Reference number:	S53729; MUID:95234757; PMID:7718624
A:Accession:	S53729
A:Status:	preliminary
A:Molecule type:	mRNA
A:Residues:	1-354, 'S', 356-461, 'D', 463-552 <GAO>
A:Cross-references:	UNIPARC:UPI00001707CF; GB:U12149; NID:9862472; PIDN:AAA85033.1; PID:
C:Function:	
A:Description:	catalyzes the formation of peptidyl-serine-phosphate or peptidyl-threonin
C:Superfamily:	SNF1-related protein kinase; protein kinase homology
C:Keywords:	ATP; magnesium; phosphotransferase; serine/threonine-specific protein kinase
F:14-268/Domain:	protein kinase homology <KIN>
F:22-30/Region:	protein kinase ATP-binding motif
F:45,64,139,141/Active site:	Lys, Glu, Asp, Lys #status predicted
F:144,148/Binding site:	magnesium (Asn, Asp) #status predicted
Query Match	18.7%; Score 644; DB 1; Length 552;
Best Local Similarity	43.3%; Pred. No. 1.7e-21;
Matches 139; Conservative	50; Mismatches 116; Indels 16; Gaps 6;
QY	10 YYELHETIGTGGFAKVKLACHILTGMVAIKMDKNTLGS--DLPRKTEIEALKNLRHQ 67
DB	15 HYLVDGTLGVGTGKVKIGEHQHTGHKVAVKILNRQKIRSLDVGKIKREIQNLKLFRRP 74
QY	68 HICQLYHVLLETANKIPMWLEFCPGGELFDYIIISQDRLSEETRVPRQIVSAVAVHSHQ 127

Db	75 HIKLQVISTPTDFPMWYVSGGELFDYICKGRVEVEARLFFQQLSAVDYCHRM 134
QY	128 YAHRLDKPENLLFDEYHKLKLDIFGLCAKPKGNKDYHLQTCGSLAYAAPELIQKSYLG 187
DB	135 VVRDLKPEVLLDAHNAKIADFGLSNMMSDGE--FLRTSCGSPNYAAPEVISGRLYAG 192
QY	188 SEADVWSMGLLYVLMCGFLPFDDNNMVALYKIMRGKYDVPKWLSPSSILLLOQMLOYD 247
DB	193 PEVDIWSGCVILYALLCGTLPPDDEHVPTLFKIRGGVFYIPEYLNRSVATLLMHMLQYD 252
QY	248 PKKRISMKNLNHPWIMQD---YNYVPEWQSKNPFTH--LDDDCVTELSVHHR-NNRQT 300
DB	253 PLKRTATKIDIREHWFQDLPVLPF-----EDPSYDANVIDEAVKEVCEKPECTESEV 307
QY	301 MEDLISLWQYDHLTATYLLLL 321
DB	308 MNSLYSGDPQDLAVAYHLII 328
RESULT 11	
T10449	probable serine/threonine-specific protein kinase (EC 2.7.1.-) - cucumber
N:Alternate names:	SNF1-related protein kinase
C:Species:	Cucumis sativus (cucumber)
C:Date:	16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 05-Oct-2004
C:Accession:	T10449
R:Gumpel, N.J.	
submitted to the EMBL Data Library,	December 1996
A:Reference number:	Z17020
A:Accession:	T10449
A:Status:	preliminary; translated from GB/EMBL/DBJ
A:Molecule type:	mRNA
A:Residues:	1-504 <GUM>
A:Cross-references:	UNIPROT:P93113; UNIPARC:UPI000004B92; EMBL:Y10036
A:Experimental source:	cv. Masterpiece; cotyledon
C:Function:	
A:Description:	catalyzes the formation of peptidyl-serine-phosphate or peptidyl-threonine
C:Superfamily:	SNF1-related protein kinase; protein kinase homology
C:Keywords:	ATP; phosphotransferase; serine/threonine-specific protein kinase
F:6-260/Domain:	protein kinase homology <KIN>
Query Match	18.6%; Score 641.5; DB 2; Length 504;
Best Local Similarity	42.2%; Pred. No. 2e-21;
Matches 146; Conservative	53; Mismatches 102; Indels 45; Gaps 9;
QY	11 YELHETIGTGGFAKVKLACHILTGMVAIKMDKNTLGS-DL-PRKTEIEALKNLRHQ 68
DB	8 YKLGKTLGLSGFGKVKIAEHALTGHKVAIKILNRKIKMLDMEEKVRRRIKILRLPMPHP 67
QY	69 ICQLYHVLLETANKIPMWLEFCPGGELFDYIIISQDRLSEETRVPRQIVSAVAVHSHQ 128
DB	68 IIRLYEVIETPSDIYVWMEYVKSGLFDYIVKGRLOEAEARNFFQIISGVVEYCHRMV 127
QY	129 AHRDLKPEVLLFDEYHKLKLDIFGLCAKPKGN--KDYH-LQTCGSLAYAAPELIQKSY 185
DB	128 VHRDLKPEVLLDSCNVCNVIADFGI-----SNIMRDGHFLKTCGSPNYAAPEVISGKLY 182
QY	186 LGSEADVWSMGLLYVLMCGFLPFDDNNMVALYKIMRGKYDVPKWLSPSSILLLOQMLO 245
DB	183 AGPEVDVWSGCVILYALLCGTLPPDDEIPLFKIKGIIYLPKSHLSSGARELIPSLMV 242
QY	246 VDPKKRISMKNLNHPWI-----MQDYNYPVWQSKNPFHLLDDCCVTEL 290
DB	243 VDPMKRITITPEIQHPWFQAHLPVYLAVPPDPMQ-----QAKK-----IDEDILQEV 290
QY	291 SVHHRNRRQTMEDLISLWQYDHLTATYLLLLAKKARGKPVRLRSS 336
DB	291 VKMGFDRNLVSLNRRIQNEATVAYLLL-----DNRFRVSS 328
RESULT 12	
T07415	probable serine/threonine-specific protein kinase (EC 2.7.1.-) PKIN1 - potato

N;Alternate names: SNF1-related protein kinase  
C;Species: Solanum tuberosum (potato)  
C;Date: 14-May-1999 #sequence\_revision 14-May-1999 #text\_change 05-Oct-2004  
C;Accession: T07415  
R;Halford, N.G.  
submitted to the EMBL Data Library, March 1996  
A;Reference number: Z16022  
A;Accession: T07415  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: DNA  
A;Residues: 1-504 <HAL>  
A;Cross-references: UNIPROT:Q41485; UNIPARC:UPI00000A0B88; EMBL:X95997; NID:g1216279; P  
C;Genetics:  
A;Gene: PKINI  
A;Introns: 62/1; 123/3; 184/3; 228/3; 290/3; 320/3; 348/3; 389/3; 466/3  
C;Function:  
A;Description: catalyzes the formation of peptidyl-serine-phosphate or peptidyl-threonine  
C;Superfamily: SNF1-related protein kinase; protein kinase homology  
C;Keywords: ATP; phosphotransferase; serine/threonine-specific protein kinase  
F;15-269/Domain: protein kinase homology <KIN>  
Query Match 18.6%; Score 640.5; DB 2; Length 504;  
Best Local Similarity 42.5%; Pred. No. 2.2e-21;  
Matches 136; Conservative 60; Mismatches 111; Indels 13; Gaps 6;  
C;QY 8 LKYYELHETIGTGGFAKVKLACHILTGEMVAIKIMDKNTLGS-DL-PRINTEIEALKNLR 65  
D; 14 LRNVGVKTLGHGSGFKVIAEHLTGKVAIKILNRKMKTPDMEELRREIKICRLFV 73  
QY 66 HQHICQLYHVLETKANKIFWVLEVCPOGELFDYIIISQDRLSEETRVVFRQIVSAVYVHS 125  
D; 74 HPVIRLYEIVETPTDIYVMYVSGELFDYIVKGRLOQEDBARKIFQOIIAGVYCHK 133  
QY 126 QGVARHDLKPENLLFDEYHKLKLIDFGLCAKPKGN--KDYH-LQTCGSLAYAAPEL 182  
D; 134 NMVVRHDLKPENLLDARRNVKIADFG-----GNIMRDGHFLKTCGSPNYAAPEVSG 198  
QY 183 KSYLGSSEADVMSGILLYLMCGFLPDDNNMVALYKIMRGKYDVPKMLSPSSILL 242  
D; 189 KLYAGPEVDVMSGVLVYALLCGTLFPDDNIPNLFKIKGIVYTLPSHLSPLARDLIPR 248  
QY 243 MLOVDPKKRISMKNLNHPHPI---MDYNYFVWQSKNPFIIHLDDCVTELSSHVHNRQ 299  
D; 249 MLIVDPKMRISVPIRQHQWFKIHLPRYLAVPDPDARQHLKCLDEISILQOVSRLMDRQ 308  
QY 300 TMDLISLWQYDHLTATYLL 319  
D; 309 LLDSLQKRIQDDATVAYLL 328  
RESULT 13  
JC1446  
serine/threonine-specific protein kinase (EC 2.7.1.-) AK21 - Arabidopsis thaliana  
N;Alternate names: protein kinase SNF1 homolog  
C;Species: Arabidopsis thaliana (mouse-ear cress)  
C;Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 05-Oct-2004  
C;Accession: JC1446; S58266; S66334  
R;LeGuen, L.; Thomas, M.; Bianchi, M.; Halford, N.G.; Kreis, M.  
Gene 120, 249-254, 1992  
A;Title: Structure and expression of a gene from Arabidopsis thaliana encoding a protein  
A;Reference number: JC1446; MUID:93013041; PMID:1339373  
A;Accession: JC1446  
A;Molecule type: DNA  
A;Residues: 1-512 <LEG>  
A;Cross-references: UNIPROT:Q38997; UNIPARC:UPI000012DE43; GB:M93023; NID:g166599; PIDN:  
R;Thuemmler, F.; Kirchner, M.; Teuber, R.; Ditttrich, P.  
submitted to the EMBL Data Library, May 1995  
A;Description: Differential accumulation of the transcripts of 22 novel protein kinase g  
A;Reference number: S58266  
A;Accession: S58266  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 144-198 <THU>

A;Cross-references: UNIPARC:UPI000009DEB0; EMBL:X86966; NID:g928909; PIDN:CAA60529.1; PII  
R;Thuemmler, F.; Kirchner, M.; Teuber, R.; Ditttrich, P.  
Plant Mol. Biol. 29, 551-565, 1995  
A;Title: Differential accumulation of the transcripts of 22 novel protein kinase genes in  
A;Reference number: S66314; MUID:96123233; PMID:8534852  
A;Accession: S66334  
A;Molecule type: DNA  
A;Residues: 144-198 <TH2>  
A;Cross-references: UNIPARC:UPI000009DEB0; EMBL:X86966; NID:g928909; PIDN:CAA60529.1; PII  
C;Comment: This enzyme plays an important role in a signal transduction cascade regulat  
C;Genetics:  
A;Gene: AKin10; AK21  
A;Introns: 64/1; 125/3; 186/3; 230/3; 292/3; 322/3; 350/3; 396/3; 475/3  
C;Function:  
A;Description: catalyzes the formation of peptidyl-serine-phosphate or peptidyl-threonine  
C;Superfamily: SNF1-related protein kinase; protein kinase homology  
C;Keywords: ATP; magnesium; phosphotransferase; serine/threonine-specific protein kinase  
F;17-271/Domain: protein kinase homology <KIN>  
F;25-33/Region: protein kinase ATP-binding motif  
F;48,67,142,144/Active site: Lys, Glu, Asp, Lys #status predicted  
F;147,151/Binding site: magnesium (Asn, Asp) #status predicted  
Query Match 18.5%; Score 637; DB 1; Length 512;  
Best Local Similarity 38.8%; Pred. No. 3.2e-21;  
Matches 151; Conservative 61; Mismatches 133; Indels 44; Gaps 11;  
C;QY 5 DELLYELHETIGTGGFAKVKLACHILTGEMVAIKIMDKNTLGS--DLPRINTEIEALK 62  
D; 13 ESILPNYKLGRTIGSGFGRVIAEHALTGKVAIKILNRKIKNMEEKVREIKILR 72  
QY 63 NLRHQHICQLYHVLETKANKIFWVLEVCPOGELFDYIIISQDRLSEETRVVFRQIVSAVAY 122  
D; 73 LFMHPHPIRLYEIVETPTDIYVMYVSGELFDYIVKGRLOQEDBARNFQQIISGVEY 132  
QY 123 VHSQGVARHDLKPENLLFDEYHKLKLIDFGLCAKPKGN--KDYH-LQTCGSLAYAAPEL 179  
D; 133 CHRNVMVVRHDLKPENLLDLSKCNKIADFG-----SNIMRDGHFLKTCGSPNYAAPEV 187  
QY 180 IQKSYLGSSEADVMSGILLYLMCGFLPDDNNMVALYKIMRGKYDVPKMLSPSSILL 239  
D; 188 ISGLYAGPEVDVMSGVLVYALLCGTLFPDDNIPNLFKIKGIVYTLPSHLSPGARDL 247  
QY 240 LQOMLOVDPKKRISMKNLNHPHPI---MDYNYFVWQSKNPFIIHLDDCVTEL 290  
D; 248 IPRMLVDPKMRISVPIRQHQWFKIHLPRYLAVPDPDTVQQA---IDEEILQEV 301  
QY 291 SVHRRNRNQTMDLISLWQYDHLTATYLLAKKARGKPVRLRLSSFSFG-----QASAT 345  
D; 302 INMGFDRNHLIESLRNTQDGTVTYYLILN-----RFRASSGYLGAEFQETWEGT 353  
QY 346 PFTDIKSNNSWLEDVTASDKNTYVAGLIDY 374  
D; 354 P-----RMHPAESVASPVSHRLPGLMEY 376  
RESULT 14  
S60304  
serine/threonine-specific protein kinase (EC 2.7.1.-) BKIN12 (version 2) - barley  
C;Species: Hordeum vulgare (barley)  
C;Date: 19-Mar-1997 #sequence\_revision 15-Aug-1997 #text\_change 05-Oct-2004  
C;Accession: S60304; S24579  
R;Halford, N.G.; Vicente-Carabajosa, J.; Sabelli, P.A.; Shewry, P.R.; Hammappel, U.; Kreis  
Plant J. 2, 791-797, 1992  
A;Title: Molecular analyses of a barley multigene family homologous to the yeast protein  
A;Reference number: S60303; MUID:93258420; PMID:1302632  
A;Accession: S60304  
A;Status: preliminary  
A;Molecule type: mRNA  
A;Residues: 1-513 <HAL>  
A;Cross-references: UNIPROT:Q40030; UNIPARC:UPI000017255D; EMBL:X65604  
R;Halford, N.G.  
submitted to the EMBL Data Library, April 1992  
A;Reference number: S24578

A;Accession: S24579  
A;Status: preliminary  
A;Molecule type: mRNA  
A;Residues: 1-61, 'A', '63-513 <HA2>  
A;Cross-references: UNIPARC:UPI00000A3F6F; EMBL: X65604; NID: g18933; PIDN: CAA46554.1; PID  
C;Function:  
A;Description: catalyzes the formation of peptidyl-serine-phosphate or peptidyl-threonin  
C;Superfamily: SNF1-related protein kinase; protein kinase homology  
C;Keywords: ATP; magnesium; phosphotransferase; serine/threonine-specific protein kinase  
F;15-272/Domain: protein kinase homology <KIN>  
F;23-31/Region: protein kinase ATP-binding motif  
F;46,65,143,145/Active site: Lys, Glu, Asp, Lys #status predicted  
F;148,152/Binding site: magnesium (Asn, Asp) #status predicted

Query Match 18.5%; Score 637; DB 1; Length 513;  
Best Local Similarity 39.4%; Pred. No. 3.2e-21;  
Matches 152; Conservative 64; Mismatches 136; Indels 34; Gaps 10;

QY 4 YDELLKYELHETIGTGFAKVKLAACHILTGEMVAIKMDKNTLGSGLPRIK--TEIEAL 61  
DB 10 HSEVLKYNLKGTLGLGTGFDVKVARNVTGQVAKILNRRKWTMEMEKGKREIKIM 69

QY 62 K---NLRHQHICOLYHLETANKIFMWLEVCYCGELFDYIIISQDRLSEEBETRVVFRQIVS 118  
DB 70 RLFDITPHPIIRVYEVIETPKDIFVVMYECNNGELLDYIIENGRLLQDEARRIFQOILA 129

QY 119 AVAYVHSQGYAHRDLKPENLLFDEYHKLKLIIDFGLCAKPKGNKDYH-LQTCGSLAYAAP 177  
DB 130 GVEYCHRMVVRDLKPENLLDSRYNVKLADEGL---SNVMDRGHFLKTSKCSLNYAAP 186

QY 178 ELIQGSKSYLGSEADVMSGILLYVLMCGFLPFDDNNVMALYKKIMRGKYDVPKWLSPSSI 237  
DB 187 EIISSKLYAGPEVDVMSGVLVALLCGSVFPDDNIPSLFRKIKGGTYLPSYLSDSAR 246

QY 238 LLQOQMLQVDPKKRISMKNLNHPWIMQDYNYPV-----EWQSKNPFHLLDDCVT 288  
DB 247 DLIPKLNIDPMKRITPHEIRVHPFKN--HLPCLYAVPPPYKEQQAQ----MIDEDILR 300

QY 289 ELSVHRRNRQTMEDLISLWQDHLTATYLLAKKARGKPVRLRLSSFCGQASATPFT 348  
DB 301 EYVNLGYDKDHVCESLWNRLQNEETVAYILL-----DDRFRSTSGYL--ADHQLM 351

QY 349 DIKSNWNL-EDVTASDKNYVAGLID 373  
DB 352 DRSFNEFTLSASASPTRNYPGIND 377

RESULT 15  
S72513  
FOG2 protein - yeast (Kluyveromyces marxianus var. lactis)  
C;Species: Kluyveromyces marxianus var. lactis, Candida sphaerica  
C;Date: 29-Jan-1998 #sequence revision 06-Feb-1998 #text change 05-Oct-2004  
C;Accession: S72513  
R;Goffrini, P.; Ficarella, A.; Donnini, C.; Lodi, T.; Puglisi, P.P.; Ferrero, I.  
Curr. Genet. 29, 316-326, 1996  
A;Title: FOG1 and FOG2 genes, required for the transcriptional activation of glucose-rep  
A;Reference number: S72513; MUID: 96171514; PMID: 8598052  
A;Accession: S72513  
A;Status: not compared with conceptual translation  
A;Molecule type: DNA  
A;Residues: 1-602 <GOF>  
A;Cross-references: UNIPROT: P87209; UNIPARC: UPI000006A88D  
A;Note: the source is designated as Kluyveromyces lactis  
C;Genetics:  
A;Gene: FOG2  
C;Function:  
A;Description: probably involved in the regulation of glucose-repressible gene expressio  
C;Superfamily: SNF1-related protein kinase; protein kinase homology  
F;33-286/Domain: protein kinase homology <KIN>

Query Match 18.5%; Score 636; DB 2; Length 602;  
Best Local Similarity 44.5%; Pred. No. 4.1e-21;  
Matches 133; Conservative 51; Mismatches 87; Indels 28; Gaps 7;

QY 11 YELHETIGTGFAKVKLAACHILTGEMVAIKMDKNTLGSGLP-RIKTEIEALKNLRHQH 68  
DB 35 YQIITLGSFGKVKLAHYHISTGQKVALKIINKKVLAKSDMQGRIEREISYLRLLRHPH 94

QY 69 ICOLYHVLLETANKIFMWLEVCYCGELFDYIIISQDRLSEEBETRVVFRQIVSAYVHSQGY 128  
DB 95 IIKLYDVIKSDEIIMVIEYA-GNELFDYIVQDRKMPQEARRFFQOIIISAVDYCHRHKI 153

QY 129 AHRDLKPENLLFDEYHKLKLIIDFGLC-AKPKGNKDYHLOTCCGSLAYAAPETIQKSYLG 187  
DB 154 VHRDLKPENLLLDHNLNVAIDFGLSNIMTDGN---FLATSCGSPNYAAPEVISGKLIYAG 210

QY 188 SEADVMSGILLYVLMCGFLPFDDNNVMALYKKIMRGKYDVPKWLSPSSITLLIQMLQVD 247  
DB 211 PEVDVMSGVLVLMCRRLPFDDDESIPVLFKNINSGVVTIPNFLSQGAASLIKKMLIVN 270

QY 248 PKKRISMKNLNHPWI---MQDYNYPVVEWQSKNPFHLLDDCVTELVSVHRRNRQTMED 303  
DB 271 PVNRIITVHEIMQDEWPKVDLPDYLPAAE-----STHQENSESKTED 311

Search completed: November 21, 2006, 14:14:15  
Job time : 47 secs

GenCore version 5.1.9  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: November 21, 2006, 14:13:40 ; Search time 53 Seconds  
(without alignments)  
1075.140 Million cell updates/sec

Title: US-10-656-598-2  
Perfect score: 3447  
Sequence: 1 MKDYDELLKYYELHETIGT.....KGDWVYKRLVEDILSSCKV 651

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 650591 seqs, 87530628 residues

Total number of hits satisfying chosen parameters: 650591

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA:\*

- 1: /EMC\_Celerra\_SID33/ptodata/2/iaa/5 COMB.pap.\*
- 2: /EMC\_Celerra\_SID33/ptodata/2/iaa/6 COMB.pap.\*
- 3: /EMC\_Celerra\_SID33/ptodata/2/iaa/7 COMB.pap.\*
- 4: /EMC\_Celerra\_SID33/ptodata/2/iaa/H COMB.pap.\*
- 5: /EMC\_Celerra\_SID33/ptodata/2/iaa/PCITUS COMB.pap.\*
- 6: /EMC\_Celerra\_SID33/ptodata/2/iaa/RE COMB.pap.\*
- 7: /EMC\_Celerra\_SID33/ptodata/2/iaa/backfiles.pap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3447	100.0	651	2	US-09-870-937-10
2	791.5	23.0	724	2	US-09-949-016-8234
3	789.5	22.9	713	2	US-09-538-092-1022
4	789.5	22.9	713	2	US-09-949-016-6214
5	788.5	22.9	729	1	US-08-677-298-2
6	788.5	22.9	729	2	US-09-523-849-33
7	770.5	22.4	691	2	US-09-949-016-8255
8	770.5	22.4	691	2	US-09-949-016-8256
9	769.5	22.3	724	2	US-09-984-890-2
10	769.5	22.3	724	2	US-10-274-194-2
11	769.5	22.3	724	2	US-10-760-407-2
12	764.5	22.2	745	2	US-09-523-849-36
13	759.5	22.0	722	2	US-09-984-890-4
14	759.5	22.0	722	2	US-10-274-194-4
15	759.5	22.0	722	2	US-10-760-407-4
16	754.5	21.9	722	2	US-08-817-832B-32
17	742.5	21.5	793	2	US-09-523-849-32
18	724.5	21.0	779	2	US-08-817-832B-31
19	669	19.4	557	2	US-09-949-016-10174
20	668.5	19.4	602	2	US-09-949-016-7417
21	668.5	19.4	602	2	US-09-949-016-7418
22	665.5	19.3	345	2	US-09-101-146-1
23	654	19.0	832	2	US-10-464-939-2
24	653	18.9	1058	2	US-10-464-939-4
25	644	18.7	552	2	US-08-557-006C-40
26	642	18.6	257	2	US-09-101-146-6

ALIGNMENTS

RESULT 1

US-09-870-937-10

; Sequence 10, Application US/09870937

; Patent No. 6974867

; GENERAL INFORMATION:

; APPLICANT: Wu, Bin

; APPLICANT: Seeley, Todd

; APPLICANT: Williams, Lewis T.

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING NEOPLASTIC DISEASE USING

; FILE OF INVENTION: CHEMOTHERAPY AND RADIATION SENSITIZERS

; FILE REFERENCE: 200130.514/PP-01623.002

; CURRENT APPLICATION NUMBER: US/09/870,937

; CURRENT FILING DATE: 2001-05-30

; NUMBER OF SEQ ID NOS: 12

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 10

; LENGTH: 651

; TYPE: PRT

; ORGANISM: Homo sapien

; US-09-870-937-10

Query Match	100.0%	Score 3447;	DB 2;	Length 651;
Best Local Similarity	100.0%	Pred. No. 2.4e-310;		
Matches 651;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	MKDYDELLKYYELHETIGTGGPAKVKLACHILTGMVAIKIMDKNTLGSDDLPRIKTEIEA	60	
DB	1	MKDYDELLKYYELHETIGTGGPAKVKLACHILTGMVAIKIMDKNTLGSDDLPRIKTEIEA	60	
QY	61	LKNTLHQHICQLYHYLETANKIFMWLEVCPCGGEDEYIISODRLSEETRVVFRQIVSAV	120	
DB	61	LKNTLHQHICQLYHYLETANKIFMWLEVCPCGGEDEYIISODRLSEETRVVFRQIVSAV	120	
QY	121	AYVHSQGVYAHRLDKPENLLFDEYHKLKLTDFGLCAKPKGNKDYHLQTCGSLAYAAPELI	180	
DB	121	AYVHSQGVYAHRLDKPENLLFDEYHKLKLTDFGLCAKPKGNKDYHLQTCGSLAYAAPELI	180	
QY	181	QKSYLGSSEADVWSMGILLVYLMCGFLPFDDNNVMAKYKIMRGKYDVPKWLSPPSILL	240	
DB	181	QKSYLGSSEADVWSMGILLVYLMCGFLPFDDNNVMAKYKIMRGKYDVPKWLSPPSILL	240	
QY	241	QOMLVDPKPKRISMKVLLNHPWIMQDYNYPVWQSKNPFHLLDDCVTELSVHHNNRQT	300	
DB	241	QOMLVDPKPKRISMKVLLNHPWIMQDYNYPVWQSKNPFHLLDDCVTELSVHHNNRQT	300	
QY	301	MEDLISLWQYDHLTATYLLLLAKKARGKPVRLRLSSFCGQASATPFTDICKNNWSLSDV	360	
DB	301	MEDLISLWQYDHLTATYLLLLAKKARGKPVRLRLSSFCGQASATPFTDICKNNWSLSDV	360	

QY 361 TASDKNYVAGLIDYDWCEDDLSTGAATPRTSQFTKYWTESNGVESKSLTPALCRTPANKL 420  
DB 361 TASDKNYVAGLIDYDWCEDDLSTGAATPRTSQFTKYWTESNGVESKSLTPALCRTPANKL 420  
QY 421 KKNENYTPKSAVKNBEYFMPEPKTPVKNQKHREILTPNRYTTPSKARNQCLKETPI 480  
DB 421 KKNENYTPKSAVKNBEYFMPEPKTPVKNQKHREILTPNRYTTPSKARNQCLKETPI 480  
QY 481 KIPVNSTGDKLMTGTVISPERCRSVELDLNOAHMEETPKRGAKEVFGSLERGLDKVIIV 540  
DB 481 KIPVNSTGDKLMTGTVISPERCRSVELDLNOAHMEETPKRGAKEVFGSLERGLDKVIIV 540  
QY 541 LTRSKRGSGARDGPRRLKHLHYNTTTRLVNPDQLNEIMSLPKKHVDVQKGYTLKCO 600  
DB 541 LTRSKRGSGARDGPRRLKHLHYNTTTRLVNPDQLNEIMSLPKKHVDVQKGYTLKCO 600  
QY 601 QSDFGKVTMQLFEVCOLOKPDVVGIRRLQRLKGDAMVYKRLVEDILSSCKV 651  
DB 601 QSDFGKVTMQLFEVCOLOKPDVVGIRRLQRLKGDAMVYKRLVEDILSSCKV 651  
RESULT 2  
US-09-949-016-8234  
; Sequence 8234, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; FILE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FASTSEQ for Windows Version 4.0  
; SEQ ID NO 8234  
; TYPE: PRT  
; ORGANISM: Human  
US-09-949-016-8234  
Query Match 23.0%; Score 791.5; DB 2; Length 724;  
Best Local Similarity 30.7%; Pred. No. 3.8e-64;  
Matches 217; Conservative 121; Mismatches 254; Indels 115; Gaps 21;  
QY 11 YELHETIGTGGFAKVKLACHILTGEMVAIKIMDKNTLG-SDLPRIKTEIEALKNLRHQHI 69  
DB 67 YRLKTKIGKNGFAKVKLARHILTGREVAIKIIDKTQLNPTSLOKLPREVIRIMKILNHPNI 126  
QY 70 COLYHLETANKIFMVLEYCPGGLFDYIISQRLSEETRVVFRQIVSAVAVHVSQGYA 129  
DB 127 VKLFEVETETKTYLIMEVAGSGEVDYLVAGRMKEKARSKEFRQIVSAVQYCHOKRIV 186  
QY 130 HRDLKPENLLFDYHKLKIDFGLCAK-PKGNKDYHLQTCGSLAYAAPELIQKGYLGS 188  
DB 187 HRDLKAENLLDADNMNIKIADFGSNEFTVGK--LDTFGSGPPYAAAPELFOGKYDGP 243  
QY 189 EADVMSGILLVLMCGFLPDDNVMALYKKIMRGKYDVPKWLSPSSILLQOQLQVDP 248  
DB 244 EVDVMSGLVLYTLVSGSLPFDGQNLKELRVRIGRKYRIPFYMSTDCENLLKRFVLNAP 303  
QY 249 KGRISMKNLNHPWIMQDYNYPVWQSKNPFIFH----LDDCCVTLSVHRRNNRQTMEDL 304  
DB 304 IKRGTEIQIMKDRWI---NAGHEEDELKPFVEPELIDISQKRIDIMVGVNGYQEBIQES 359  
QY 305 ISLWQYDHLTATYLLILLAKARQKPVRLRLSSFCGQASATPTFDI-----KSNWWS 356  
DB 360 LSKMKYDEITATYLLLLGRKSSVRP-----SSDLNNSGTQSPHHKVRQSVSSSKORRYS 414

QY 357 -----LEDVTASDKNYVAGLIDYDWCEDDLST-----GAATP- 388  
DB 415 DHAGPAIFSVVAYPKRSQTSADSDLKEDGISSRKSSGSGVGGKGIAPASPMLGNASPN 474  
QY 389 -----RTSQFTKYWTESNGVESKSLTPALCRTPANKLKNKENVYTPKSAVKN-BEYF 439  
DB 475 KADIPERKKSSTVPSSNTASGWMTRNRYTVGCSERTADR-----HSVIOQNGKENS 524  
QY 440 MPEBKTPVKNQKHREILTPNRYTTPSKARNQCLKETPIKIPVNSTGDKLMTGVIIV 499  
DB 525 TIPDORTPV-ASTHISIAAATPDR-----IRFP--RGTASRSTFHQGP 564  
QY 500 -ERRC-----RSVELDLNOAHMEETPKRGAKEVFGSLERGLDKVIIVLTRSK-----R 546  
DB 565 RERRATYNGPPASPSSLSHEATPLSQTSRSGTNLF-----SKLTSKLTSRNVSAEQ 617  
QY 547 KGSARDG-PRRLKHLHYNTTTRLVNPDQLNEIMSLPKKHVDVQK-GYTLKCOQSDP 604  
DB 618 KDNKEAKPRSLRFTWSMKTTSSMDPGDMRREIRKVLDAANNCDYEQRERFLLFCVHGDGH 677  
QY 605 GKVTMQLFEVCOLOKPDVVGIRRLQRLKGDAMVYKRLVEDILSSCKV 651  
DB 678 AENLVQWEMEVCCKPRLSLNGVRFRKRSIGTSAFKNIASKIANELKL 724  
RESULT 3  
US-09-538-092-1022  
; Sequence 1022, Application US/09538092  
; Patent No. 6753314  
; GENERAL INFORMATION:  
; APPLICANT: Giot, Loic  
; APPLICANT: Mansfield, Traci A.  
; TITLE OF INVENTION: Protein-Protein Complexes and Method of Using Same  
; FILE REFERENCE: 15966-542  
; CURRENT APPLICATION NUMBER: US/09/538,092  
; CURRENT FILING DATE: 2000-03-29  
; PRIOR APPLICATION NUMBER: 60/127,352  
; PRIOR FILING DATE: 1999-04-01  
; PRIOR APPLICATION NUMBER: 60/178,965  
; PRIOR FILING DATE: 2000-02-01  
; NUMBER OF SEQ ID NOS: 1387  
; SOFTWARE: CurPatSeqformatter Version 0.9  
; SEQ ID NO 1022  
; LENGTH: 713  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc.feature  
; LOCATION: (0)..(0)  
; OTHER INFORMATION: Polypeptide Accession Number P27448  
US-09-538-092-1022  
Query Match 22.9%; Score 789.5; DB 2; Length 713;  
Best Local Similarity 30.6%; Pred. No. 5.6e-64;  
Matches 217; Conservative 119; Mismatches 254; Indels 119; Gaps 22;  
QY 11 YELHETIGTGGFAKVKLACHILTGEMVAIKIMDKNTLG-SDLPRIKTEIEALKNLRHQHI 69  
DB 56 YRLKTKIGKNGFAKVKLARHILTGREVAIKIIDKTQLNPTSLOKLPREVIRIMKILNHPNI 115  
QY 70 COLYHLETANKIFMVLEYCPGGLFDYIISQRLSEETRVVFRQIVSAVAVHVSQGYA 129  
DB 116 VKLFEVETQKTYLIMEVAGSGKVDYLVAGRMKEKARSKEFRQIVSAVQYCHOKRIV 175  
QY 130 HRDLKPENLLFDEYHKLKIDFGLCAK-PKGNKDYHLQTCGSLAYAAPELIQKGYLGS 188  
DB 176 HRDLKAENLLDADNMNIKIADFGSNEFTVGK--LDTFGSGPPYAAAPELFOGKYDGP 232  
QY 189 EADVMSGILLVLMCGFLPDDNVMALYKKIMRGKYDVPKWLSPSSILLQOQLQVDP 248  
DB 233 EVDVMSGLVLYTLVSGSLPFDGQNLKELRVRIGRKYRIPFYMSTDCENLLKRFVLNAP 292





TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-677-298-2

Query Match 22.9%; Score 788.5; DB 1; Length 729;  
Best Local Similarity 30.0%; Pred. No. 7.2e-64;  
Matches 216; Conservative 124; Mismatches 256; Indels 123; Gaps 20;

```
QY 11 YELHETIGTGGFAKVKLACHILTGEMVAIKIMDKNTLG-SDLPRIKTEIEALKNLRHQHI 69
DB 56 YRLKKTIGKGNFAKVKLARHILTGREVAIKIDKLTQNLPTSLQKLFREVRIMKILNHPNI 115
QY 70 COLYHVLETANKIFMWLEYPGCGELFDYIISODRLSEEBETRVVFRQIVSAVAVHSQGYA 129
DB 116 VKLFEVIEKTEKTYLIMEYASGGEVFDYLVAGHMKKEARSKFRQIVSAVQYCHQKRV 175
QY 130 HRDLKPEMLLFDEYHKLKIDFGCAK-PKGNKDYHLQTCGSLAYAAPELIQGSYLG 188
DB 176 HRDLKAENLLDADNMNIKIADFGSNEFTVGK---LDTFCGSPPPYAAPELFGQKKYDGP 232
QY 189 EADVWSGILLVLMCGFLPDDDDNMALYKIMRGKYDVPKWLSPSSILLQQMLQVDP 248
DB 233 EVDVWSGLVILYTLVSGSLPFDGQNLKELRERVLRGKIRIPFYMSTDCENLLKRFVLNLP 292
QY 249 KKRISMKNLLNHPWIMQDYNYPVEMQSKNPFTH---LDDDCVTELSVHHRNNRQTMEDL 304
DB 293 IKRGTELEQIMKORNI---NAGHEEDELKPFVEPELDISDQKREIDIMVGMGYSQEBIQES 348
QY 305 ISLMQYDHLTATYLLLLAKKARGKVPRLRLSPSCQASATPFTDIKSNW----- 355
DB 349 LSKMKYDEITATY-LLLGKSKSELDASDSSSSNLSLAKVRPSSDLNNSGTQSPHHKVQR 407
QY 356 -----SLEDVTASDKNYVAGLIDYDWCEDDLST----- 383
DB 408 SVSSSQKORRYSDHAGPAIPSVVAYPKRSQTSADGDLKEDGISRKSSGSAVGGKGIAP 467
QY 384 -----GAATP-----RTSQFTKYWTESGVESKSLTPALCRTPANKLNKNENVT 428
DB 468 ASPMLGNASNPNKADIPERKKSSTVPSSNTASGGMTRNTYVCSERTTADR----- 518
QY 429 PKSAVKN-BEYFMPPEKTPVKNQKHREILTPNRYTTPSKARNOCLEKETPIKIPVNST 487
DB 519 -HSVIONGENSTIPDQTPV-ASTHSISSAATPDR-----IRFP---R 557
QY 488 GTDKLMTGVIISP-ERRC-----RSVELDLNQAHEETPKRGAKVFGSLERGLDKVIT 539
DB 558 GTASRSTFHGQPRERTATYNGPPSPSLSHSATPLSOTRSRGSTNLF-----SKLTS 610
QY 540 VLTRSK-----RKGSDRG-PRRLKLYNVTTTLLVNPDLNLEIMSLPKKHVDVQK- 592
DB 611 KLTRSRNVSAEQDENKEAKPRSLFTWSMKTSSMDPGDMREIRKVLNANNCDEQRE 670
QY 593 GYTLKQOTSDFGKVTMQFLEVCQKQPDVVGIRQRKLGDAWYKRLVEDILSSCKV 651
DB 671 RPLLCVHGDAENLVQWEMEVCKLPRLSLNGVRFKRSIGTSIAFNKIASIANELKL 729
```

RESULT 6

US-09-523-849-33  
Sequence 33, Application US/09523849  
Patent No. 6458561  
GENERAL INFORMATION:  
APPLICANT: Bandman, Olga  
APPLICANT: Molteni, Angela  
APPLICANT: Magnaghi, Paola  
APPLICANT: Bosotti, Roberta  
APPLICANT: Scacheri, Emanuela  
APPLICANT: Isacchi, Antonella  
APPLICANT: Hodgson, Dave  
TITLE OF INVENTION: HUMAN NIMI KINASE  
FILE REFERENCE: PC-0009 US  
CURRENT APPLICATION NUMBER: US/09/523,849

CURRENT FILING DATE: 2000-03-13  
NUMBER OF SEQ ID NOS: 39  
SOFTWARE: PERL Program  
SEQ ID NO 33  
LENGTH: 729  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: misc feature  
OTHER INFORMATION: GenBank Accession No. 6458561 g3089349  
US-09-523-849-33

Query Match 22.9%; Score 788.5; DB 2; Length 729;  
Best Local Similarity 30.0%; Pred. No. 7.2e-64;  
Matches 216; Conservative 124; Mismatches 256; Indels 123; Gaps 20;

```
QY 11 YELHETIGTGGFAKVKLACHILTGEMVAIKIMDKNTLG-SDLPRIKTEIEALKNLRHQHI 69
DB 56 YRLKKTIGKGNFAKVKLARHILTGREVAIKIDKLTQNLPTSLQKLFREVRIMKILNHPNI 115
QY 70 COLYHVLETANKIFMWLEYPGCGELFDYIISODRLSEEBETRVVFRQIVSAVAVHSQGYA 129
DB 116 VKLFEVIEKTEKTYLIMEYASGGEVFDYLVAGHMKKEARSKFRQIVSAVQYCHQKRV 175
QY 130 HRDLKPEMLLFDEYHKLKIDFGCAK-PKGNKDYHLQTCGSLAYAAPELIQGSYLG 188
DB 176 HRDLKAENLLDADNMNIKIADFGSNEFTVGK---LDTFCGSPPPYAAPELFGQKKYDGP 232
QY 189 EADVWSGILLVLMCGFLPDDDDNMALYKIMRGKYDVPKWLSPSSILLQQMLQVDP 248
DB 233 EVDVWSGLVILYTLVSGSLPFDGQNLKELRERVLRGKIRIPFYMSTDCENLLKRFVLNLP 292
QY 249 KKRISMKNLLNHPWIMQDYNYPVEMQSKNPFTH---LDDDCVTELSVHHRNNRQTMEDL 304
DB 293 IKRGTELEQIMKORNI---NAGHEEDELKPFVEPELDISDQKREIDIMVGMGYSQEBIQES 348
QY 305 ISLMQYDHLTATYLLLLAKKARGKVPRLRLSPSCQASATPFTDIKSNW----- 355
DB 349 LSKMKYDEITATY-LLLGKSKSELDASDSSSSNLSLAKVRPSSDLNNSGTQSPHHKVQR 407
QY 356 -----SLEDVTASDKNYVAGLIDYDWCEDDLST----- 383
DB 408 SVSSSQKORRYSDHAGPAIPSVVAYPKRSQTSADGDLKEDGISRKSSGSAVGGKGIAP 467
QY 384 -----GAATP-----RTSQFTKYWTESGVESKSLTPALCRTPANKLNKNENVT 428
DB 468 ASPMLGNASNPNKADIPERKKSSTVPSSNTASGGMTRNTYVCSERTTADR----- 518
QY 429 PKSAVKN-BEYFMPPEKTPVKNQKHREILTPNRYTTPSKARNOCLEKETPIKIPVNST 487
DB 519 -HSVIONGENSTIPDQTPV-ASTHSISSAATPDR-----IRFP---R 557
QY 488 GTDKLMTGVIISP-ERRC-----RSVELDLNQAHEETPKRGAKVFGSLERGLDKVIT 539
DB 558 GTASRSTFHGQPRERTATYNGPPSPSLSHSATPLSOTRSRGSTNLF-----SKLTS 610
QY 540 VLTRSK-----RKGSDRG-PRRLKLYNVTTTLLVNPDLNLEIMSLPKKHVDVQK- 592
DB 611 KLTRSRNVSAEQDENKEAKPRSLFTWSMKTSSMDPGDMREIRKVLNANNCDEQRE 670
QY 593 GYTLKQOTSDFGKVTMQFLEVCQKQPDVVGIRQRKLGDAWYKRLVEDILSSCKV 651
DB 671 RPLLCVHGDAENLVQWEMEVCKLPRLSLNGVRFKRSIGTSIAFNKIASIANELKL 729
```

RESULT 7

US-09-949-016-8255  
Sequence 8255, Application US/09949016  
Patent No. 6812339  
GENERAL INFORMATION:  
APPLICANT: VENTER, J. Craig et al.  
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF

```
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8255
; LENGTH: 691
; TYPE: PRT
; ORGANISM: Human
; US-09-949-016-8255

Query Match      22.4%; Score 770.5; DB 2; Length 691;
Best Local Similarity 30.5%; Pred. No. 3.1e-62;
Matches 219; Conservative 123; Mismatches 252; Indels 125; Gaps 24;

.QY 11 YELHETIGTGGAFAKVLACHILITGEMVAIKIMDKNTL-GSDLPRIKTEIEALKNLRHQHI 69
Db 20 YRLKKTIGKGNFAKVLARHILITGKEVAVKIIDKTQLNSSSLQKLFREVRIMKVLNHPNI 79
.QY 70 CQLYHVLETANKIFMWLEYCPGGLFDYIISODRLSEBETRVVFRQIVSAVAVHSGGYA 129
Db 80 VKLFEVETETKLYLMEVYASGGEVDFYLVHGRMEKEARAKFRQIVSAVQYCHQKFTV 139
.QY 130 HRDLKPNLLFDYHKLIDFGLCAK-PKGNKDYHLOTCCGSLAAYAAPELQKSYLGS 188
Db 140 HRDLKAENLLDADNMKIADFGNEFTFGNK---LDTFCGSPPYAAPELFGQKYDGP 196
.QY 189 EADVSMGILLYVLMCGFLPFDDNNMALKYKIMRGKYDVPKWLSPSSILLQOQLQVDP 248
Db 197 EVDVNSLGVILYLVSGSLPFDQNLKELRERVLRGKYRIPFYMSDTCENLLKFLILNP 256
.QY 249 KGRISMKNLNHPIMQDYNYPVWOSKNPFIH---LDDDCVTELSVHHRNNRQTMEDL 304
Db 257 SKRGTELEQIMKDRWM---NVGHEDDELKPYVEPLDYKDPRTTELMSVSMGYTREBIQDS 312
.QY 305 ISLWQYDHLTATYLLALLAKA--RQKPVRL--RLSSFSCGOASATPTTDIKNNWSLEVD 360
Db 313 LVGQRYNEVMATYLLGYKSSLEGGDTITLKPRPSADLTNSSAPSPSHKVQ-----RSV 366
.QY 361 TASDKNYVAGLIDYDWCEDDLSTGAATPRTSOFTKYWTSNGVESKSLTPALCRTPANKL 420
Db 367 SANPKQR-----RPSDQAGPAIPTSNYSKK-TQSNNAENK--RPEEDRESGRKA 413
.QY 421 KKNENYTPKSAVKNEEYFMFPPEKTPVKNQKHREILTPN-----RYTTPSKARNQC 474
Db 414 SSTAKV-----PASPLPGLERKK---TTPTSTNSVLSTSTNRSRNP 453
.QY 475 LKE-----TPIKIPVNSTGTDLMTGVISPER-----R 502
Db 454 LLERASLQASIQNGKSTAPQRPVAPSPSAHNISSSGGADPTNPPRGVSSRSTFHAGQ 513
.QY 503 CRSVELDLQAQHEETPK-----RKA--KVFGS-----LERGL-----DKVITVL 541
Db 514 LRQVR-DQQLPYGVTPASPSGHSQGRGASGIFSKFTSKFVRRNLNPEPSKDRVETLR 572
.QY 542 T-----RSKRGSGARDG-PRRLKLYNVVTRVLPDQLNEIMSILPKKHVDF-VOK 592
Db 573 PHVVGSGNDKEKEEFREAKPRSLRFTWSMKTTSMEPEMREIRKVLNDANSQCSSELHE 632
.QY 593 GYTLKQCTQSDFGKVTMQFEVLCQKPDVVGIRQRLKGDWVYKRLVEDILSSCKV 651
Db 633 KYMLLCWHGTGPHEDFVQWEMEYCKLPRLSLNGVRFKRISGTSMAFKNIASKIANELKL 691
```

RESULT 8

US-09-949-016-8256

; Sequence 8256, Application US/09949016

```
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8256
; LENGTH: 691
; TYPE: PRT
; ORGANISM: Human
; US-09-949-016-8256

Query Match      22.4%; Score 770.5; DB 2; Length 691;
Best Local Similarity 30.5%; Pred. No. 3.1e-62;
Matches 219; Conservative 123; Mismatches 252; Indels 125; Gaps 24;

.QY 11 YELHETIGTGGAFAKVLACHILITGEMVAIKIMDKNTL-GSDLPRIKTEIEALKNLRHQHI 69
Db 20 YRLKKTIGKGNFAKVLARHILITGKEVAVKIIDKTQLNSSSLQKLFREVRIMKVLNHPNI 79
.QY 70 CQLYHVLETANKIFMWLEYCPGGLFDYIISODRLSEBETRVVFRQIVSAVAVHSGGYA 129
Db 80 VKLFEVETETKLYLMEVYASGGEVDFYLVHGRMEKEARAKFRQIVSAVQYCHQKFTV 139
.QY 130 HRDLKPNLLFDYHKLIDFGLCAK-PKGNKDYHLOTCCGSLAAYAAPELQKSYLGS 188
Db 140 HRDLKAENLLDADNMKIADFGNEFTFGNK---LDTFCGSPPYAAPELFGQKYDGP 196
.QY 189 EADVSMGILLYVLMCGFLPFDDNNMALKYKIMRGKYDVPKWLSPSSILLQOQLQVDP 248
Db 197 EVDVNSLGVILYLVSGSLPFDQNLKELRERVLRGKYRIPFYMSDTCENLLKFLILNP 256
.QY 249 KGRISMKNLNHPIMQDYNYPVWOSKNPFIH---LDDDCVTELSVHHRNNRQTMEDL 304
Db 257 SKRGTELEQIMKDRWM---NVGHEDDELKPYVEPLDYKDPRTTELMSVSMGYTREBIQDS 312
.QY 305 ISLWQYDHLTATYLLALLAKA--RQKPVRL--RLSSFSCGOASATPTTDIKNNWSLEVD 360
Db 313 LVGQRYNEVMATYLLGYKSSLEGGDTITLKPRPSADLTNSSAPSPSHKVQ-----RSV 366
.QY 361 TASDKNYVAGLIDYDWCEDDLSTGAATPRTSOFTKYWTSNGVESKSLTPALCRTPANKL 420
Db 367 SANPKQR-----RPSDQAGPAIPTSNYSKK-TQSNNAENK--RPEEDRESGRKA 413
.QY 421 KKNENYTPKSAVKNEEYFMFPPEKTPVKNQKHREILTPN-----RYTTPSKARNQC 474
Db 414 SSTAKV-----PASPLPGLERKK---TTPTSTNSVLSTSTNRSRNP 453
.QY 475 LKE-----TPIKIPVNSTGTDLMTGVISPER-----R 502
Db 454 LLERASLQASIQNGKSTAPQRPVAPSPSAHNISSSGGADPTNPPRGVSSRSTFHAGQ 513
.QY 503 CRSVELDLQAQHEETPK-----RKA--KVFGS-----LERGL-----DKVITVL 541
Db 514 LRQVR-DQQLPYGVTPASPSGHSQGRGASGIFSKFTSKFVRRNLNPEPSKDRVETLR 572
.QY 542 T-----RSKRGSGARDG-PRRLKLYNVVTRVLPDQLNEIMSILPKKHVDF-VOK 592
Db 573 PHVVGSGNDKEKEEFREAKPRSLRFTWSMKTTSMEPEMREIRKVLNDANSQCSSELHE 632
.QY 593 GYTLKQCTQSDFGKVTMQFEVLCQKPDVVGIRQRLKGDWVYKRLVEDILSSCKV 651
Db 633 KYMLLCWHGTGPHEDFVQWEMEYCKLPRLSLNGVRFKRISGTSMAFKNIASKIANELKL 691
```

RESULT 9  
US-09-984-890-2  
; Sequence 2, Application US/09984890  
; Patent No. 6492156  
; GENERAL INFORMATION:  
; APPLICANT: YAN, Chunhua et al.  
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC  
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES  
; TITLE OF INVENTION: THEREOF  
; FILE REFERENCE: CL001306  
; CURRENT APPLICATION NUMBER: US/09/984,890  
; CURRENT FILING DATE: 2001-10-31  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 2  
; LENGTH: 724  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-984-890-2

Query Match 22.3%; Score 769.5; DB 2; Length 724;  
Best Local Similarity 30.3%; Pred. No. 4.1e-62;  
Matches 218; Conservative 124; Mismatches 252; Indels 125; Gaps 24;  
  
Qy 11 YELHETIGTGGAFAKVLACHILTGEMVAIKIMDKNTL-GSDLPRIKTEIEALKNLRHQHI 69  
Db 53 YRLKKTIGKGNFAKVLARHILTGKEVAVKIIDKTQLNSSLQKLFREVRIMKVLNHPNI 112  
  
Qy 70 COLYHVLETANKIFMWLEYPGCGELFDYIISQRLSEETRVVFRQIVSAVAVVHSSQGYA 129  
Db 113 VKLFEVIEETKTYLVMEYASGEVFDYLVAGHMKKEARAKFRQVVSAYVQYCHQKFIV 172  
  
Qy 130 HRDLKPENLLFDYHKLKLIIDFGLCAK-PKGNKYHLQTCGSLAYAAPELIQGSYLS 188  
Db 173 HRDLKAENLLLDADMNKIADFGSNEFTFGNK---LDTFCGSPPYAAPELFGQKKYDGP 229  
  
Qy 189 EADVMSGILLVLMCGFLPFDDNNMALKYKIMRGKYDVPKWLSPSSILLQOQLQVDP 248  
Db 230 EVDVMSGLVLYLVSGSLPFDDQNLKELREVRVLRGKYRIPFYMTSDCENLKKFILLNP 289  
  
Qy 249 KKRISMKNLLNHPWIMQDYNYPVEWQSKNPFIIH---LDDCCVTELSVHRRNNRQTMEDL 304  
Db 290 SKRGTLQIMKDRWM---NVGHEDDELKPYVEPLDYKDPRTTELWVSMGYTREIIQDS 345  
  
Qy 305 ISLMQYDHLTATYLLLLAKKA--RGKPVRL--RLSSFSCGQASATPFTDIKSNWLSLEVD 360  
Db 346 LVGQRYNEVMATYLLLYGKSSLEGGDTITLKPRAADLTNSSAPSPSHKVQ-----RSV 399  
  
Qy 361 TASDKNYVAGLIDYDWCEDDLSTGAATPRTSQFTKYWTESNGVESKSLTPALCRTPANKL 420  
Db 400 SANPKOR-----RFSDAQGAIPAITSNSYSKK-TOSNNAENK--RPEEDRESGRKA 446  
  
Qy 421 KKNENYVTPKSAVKNEEYFMFPPEKTPVNKNQHKREILTPN-----RYTTPSKARNOC 474  
Db 447 SSTAKV-----PASPLPGLERKK---TTPTSTNSVLSTSTNSRNSP 486  
  
Qy 475 LKE-----TPIKIPVNSTGTDKLMTGVISPER-----R 502  
Db 487 LLERASLQASIQNGKDSAPQRPVVPASPSAHNIISSGGAPDRTNPPRGVSSRSTFHAQ 546  
  
Qy 503 CRSVELDLNQAHEETPK-----RKA--KVFGS-----LERGL-----DKVITVL 541  
Db 547 LRQVR-DQONLPYGVTPASPSGHSQRRGASGIFSFTSKFVRRNLNBPESKORVETLR 605  
  
Qy 542 T-----RSKRKGSARDG-PRRLKLYNNVTTRLVNPDQLNEMSILPKKHVDF-VOK 592  
Db 606 PHVVGSGNDKEKEEFREAKPRSLRTWSMKTTSMEPEMNMREIRKVLNDANSQCSELHE 665  
  
Qy 593 GYTLKQCTQSDFGKVTMQFELVCQIQKPDVVGIRQRLLKGDWVYKRLVEDILSSCKV 651  
Db 666 KYMLLCMHGTGPHGDFVQWEMEVCKLPRLSLNGVRFKRIISGTSMAFPKNIASKIANELKL 724

RESULT 10  
US-10-274-194-2  
; Sequence 2, Application US/10274194  
; Patent No. 6706511  
; GENERAL INFORMATION:  
; APPLICANT: YAN, Chunhua et al.  
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC  
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES  
; TITLE OF INVENTION: THEREOF  
; FILE REFERENCE: CL001306DIV  
; CURRENT APPLICATION NUMBER: US/10/274,194  
; CURRENT FILING DATE: 2003-12-18  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 2  
; LENGTH: 724  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-274-194-2

Query Match 22.3%; Score 769.5; DB 2; Length 724;  
Best Local Similarity 30.3%; Pred. No. 4.1e-62;  
Matches 218; Conservative 124; Mismatches 252; Indels 125; Gaps 24;  
  
Qy 11 YELHETIGTGGAFAKVLACHILTGEMVAIKIMDKNTL-GSDLPRIKTEIEALKNLRHQHI 69  
Db 53 YRLKKTIGKGNFAKVLARHILTGKEVAVKIIDKTQLNSSLQKLFREVRIMKVLNHPNI 112  
  
Qy 70 COLYHVLETANKIFMWLEYPGCGELFDYIISQRLSEETRVVFRQIVSAVAVVHSSQGYA 129  
Db 113 VKLFEVIEETKTYLVMEYASGEVFDYLVAGHMKKEARAKFRQVVSAYVQYCHQKFIV 172  
  
Qy 130 HRDLKPENLLFDYHKLKLIIDFGLCAK-PKGNKYHLQTCGSLAYAAPELIQGSYLS 188  
Db 173 HRDLKAENLLLDADMNKIADFGSNEFTFGNK---LDTFCGSPPYAAPELFGQKKYDGP 229  
  
Qy 189 EADVMSGILLVLMCGFLPFDDNNMALKYKIMRGKYDVPKWLSPSSILLQOQLQVDP 248  
Db 230 EVDVMSGLVLYLVSGSLPFDDQNLKELREVRVLRGKYRIPFYMTSDCENLKKFILLNP 289  
  
Qy 249 KKRISMKNLLNHPWIMQDYNYPVEWQSKNPFIIH---LDDCCVTELSVHRRNNRQTMEDL 304  
Db 290 SKRGTLQIMKDRWM---NVGHEDDELKPYVEPLDYKDPRTTELWVSMGYTREIIQDS 345  
  
Qy 305 ISLMQYDHLTATYLLLLAKKA--RGKPVRL--RLSSFSCGQASATPFTDIKSNWLSLEVD 360  
Db 346 LVGQRYNEVMATYLLLYGKSSLEGGDTITLKPRAADLTNSSAPSPSHKVQ-----RSV 399  
  
Qy 361 TASDKNYVAGLIDYDWCEDDLSTGAATPRTSQFTKYWTESNGVESKSLTPALCRTPANKL 420  
Db 400 SANPKOR-----RFSDAQGAIPAITSNSYSKK-TOSNNAENK--RPEEDRESGRKA 446  
  
Qy 421 KKNENYVTPKSAVKNEEYFMFPPEKTPVNKNQHKREILTPN-----RYTTPSKARNOC 474  
Db 447 SSTAKV-----PASPLPGLERKK---TTPTSTNSVLSTSTNSRNSP 486  
  
Qy 475 LKE-----TPIKIPVNSTGTDKLMTGVISPER-----R 502  
Db 487 LLERASLQASIQNGKDSAPQRPVVPASPSAHNIISSGGAPDRTNPPRGVSSRSTFHAQ 546  
  
Qy 503 CRSVELDLNQAHEETPK-----RKA--KVFGS-----LERGL-----DKVITVL 541  
Db 547 LRQVR-DQONLPYGVTPASPSGHSQRRGASGIFSFTSKFVRRNLNBPESKORVETLR 605  
  
Qy 542 T-----RSKRKGSARDG-PRRLKLYNNVTTRLVNPDQLNEMSILPKKHVDF-VOK 592  
Db 606 PHVVGSGNDKEKEEFREAKPRSLRTWSMKTTSMEPEMNMREIRKVLNDANSQCSELHE 665  
  
Qy 593 GYTLKQCTQSDFGKVTMQFELVCQIQKPDVVGIRQRLLKGDWVYKRLVEDILSSCKV 651  
Db 666 KYMLLCMHGTGPHGDFVQWEMEVCKLPRLSLNGVRFKRIISGTSMAFPKNIASKIANELKL 724

RESULT 11  
US-10-760-407-2  
; Sequence 2, Application US/10760407  
; Patent No. 6930173  
; GENERAL INFORMATION:  
; APPLICANT: YAN, Chunhua et al.  
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC  
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES  
; TITLE OF INVENTION: THEREOF  
; FILE REFERENCE: CL001306-DIV II  
; CURRENT APPLICATION NUMBER: US/10760,407  
; CURRENT FILING DATE: 2004-01-21  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 2  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-760-407-2

Query Match 22.3%; Score 769.5; DB 2; Length 724;  
Best Local Similarity 30.3%; Pred. No. 4.1e-62;  
Matches 218; Conservative 124; Mismatches 252; Indels 125; Gaps 24;

Qy 11 YELHETIGTGGFAKVKLACHILTGEMVAIKIMDKNTL-GSDLPRIKTEIALKNLRHQHI 69  
Db 53 YRLKTIKGNFAKVKLARHILTGKEVAVKIIDKTLNSSLQKLPREVIMKVLNHPNI 112

Qy 70 COLYHVLETANKIFMWLEPCGGELFDYIIISODRLSEETRVVFRQIVSAVAVHSGGYA 129  
Db 113 VKLFEVETETKTLVLMVYASGGEVFDYLVAGHMKKEARAKFRQIVSAVQYCHOKFIV 172

Qy 130 HRDLKPNLLFDEYHKLKIDFGLCAK-PKGNKDYHLQTCGSLAYAAPELIGKSVLGS 188  
Db 173 HRDLKAENLLDADNMNIADFGSNEFTFGNK---LDTFCGSPPYAAPELFGKGYDGP 229

Qy 189 EADVMSGILLVYLMCGFLPDDDDNMALYKIMRGKYDVPKWLSPSSIIILLQOMLQVDP 248  
Db 230 EVDVMSGLVILYLVSGSLPFDGQNLKELRERVLRGKYRPFYFMTDNCNLLKFLILNP 289

Qy 249 KKRISMKNLNHPWIMQDYNYPVWQSKNPFIIH---LDDCVTELSTVHRRNNRQTMEDL 304  
Db 290 SKRGTLLEQIMKDRWM---NVGHEDDELKPYVEPLDYKDPRTTELMSVSGYTRBEIQDS 345

Qy 305 ISLWQYDHLTATYLLLLAKKA--RGKPVRL--RLSSFCSCQASATPTPTDIKSNWSLEDV 360  
Db 346 LVQQRNEVMATYLLGYKSSLEGGDTITLKPRPSADLTNSSAPSPSHKQVQ-----RSV 399

Qy 361 TASDKNVVAGLIDYDWCEDDLSTGAATPRTSQTQKYWTESNGVESKSLTPALCRTPANKL 420  
Db 400 SANPKQR-----RPSDQAGPAIPTSNYSYK-KTQSNNAENK--RPEEDRESGRKA 446

Qy 421 KKNENYTPKSAVNEEYFMPPKPTKPNKQKHREILTPN-----RYTPPSKARNQC 474  
Db 447 SSTAKV-----PASPLPLGLERK---TTPSTNSVLSTSTNRSRNSP 486

Qy 475 LKX-----TPIKIPVNSTGDTKMTGTVISPER-----R 502  
Db 487 LLERASLGQASIONGKDSAPQVAVSPSAHNISGGGAPDRTNPRGVSSRSTFHAGQ 546

Qy 503 CRSVELDLNQAHEETPK-----RKGA--KVFSGS-----LBERGL-----DKVITVL 541  
Db 547 LRQVR-DQQNLPGVTPASPSGHSQGRGASGSIKFTSKFVRRNLNPESEKRVETLR 605

Qy 542 T-----RSKRGSGARDG--PRRLKLVNVTTLNPNOLLNEIMSILPKKHVDF-VOK 592  
Db 606 PHVVGSGGNDKEKEBEFREAKPRSLRFTWSMKTSTSSMEPNEMREIRKVLVDANSQCSLHE 665

Qy 593 GYTLKCTQSDFGKVTWQFEVLEQOLQKPDVVGIRQRLKGDWVYKRLVEDILLSCVK 651  
Db 666 KYMLLCMHGTGPHGHEFVQWEMEVCKLPRUSLNGVRPKRISGTSMAFKNIASKIANELKL 724

RESULT 12  
US-09-523-849-36  
; Sequence 36, Application US/09523849  
; Patent No. 6458561  
; GENERAL INFORMATION:  
; APPLICANT: Bandman, Olga  
; APPLICANT: Molteni, Angela  
; APPLICANT: Magnaghi, Paola  
; APPLICANT: Bosotti, Roberta  
; APPLICANT: Scaccheri, Emanuela  
; APPLICANT: Isacchi, Antonella  
; APPLICANT: Hodgson, Dave  
; TITLE OF INVENTION: HUMAN NIM1 KINASE  
; FILE REFERENCE: PC-0009 US  
; CURRENT APPLICATION NUMBER: US/09/523,849  
; CURRENT FILING DATE: 2000-03-13  
; NUMBER OF SEQ ID NOS: 39  
; SOFTWARE: PERL Program  
; SEQ ID NO 36  
; LENGTH: 745  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc feature  
; OTHER INFORMATION: GenBank Accession No. 6458561 g1749794  
US-09-523-849-36

Query Match 22.2%; Score 764.5; DB 2; Length 745;  
Best Local Similarity 29.6%; Pred. No. 1.3e-61;  
Matches 220; Conservative 136; Mismatches 267; Indels 121; Gaps 27;

Qy 11 YELHETIGTGGFAKVKLACHILTGEMVAIKIMDKNTL-GSDLPRIKTEIALKNLRHQHI 69  
Db 20 YRLKTIKGNFAKVKLARHILTGKEVAVKIIDKTLNSSLQKLPREVIMKVLNHPNI 79

Qy 70 COLYHVLETANKIFMWLEPCGGELFDYIIISODRLSEETRVVFRQIVSAVAVHSGGYA 129  
Db 80 VKLFEVETETKTLVLMVYASGGEVFDYLVAGHMKKEARAKFRQIVSAVQYCHOKFIV 139

Qy 130 HRDLKPNLLFDEYHKLKIDFGLCAK-PKGNKDYHLQTCGSLAYAAPELIGKSVLGS 188  
Db 140 HRDLKAENLLDADNMNIADFGSNEFTFGNK---LDTFCGSPPYAAPELFGKGYDGP 196

Qy 189 EADVMSGILLVYLMCGFLPDDDDNMALYKIMRGKYDVPKWLSPSSIIILLQOMLQVDP 248  
Db 197 EVDVMSGLVILYLVSGSLPFDGQNLKELRERVLRGKYRPFYFMTDNCNLLKFLILNP 256

Qy 249 KKRISMKNLNHPWIMQDYNYPVWQSKNPFIIH---LDDCVTELSTVHRRNNRQTMEDL 304  
Db 257 SKRGTLLEQIMKDRWM---NVGHEDDELKPYVEPLDYKDPRTTELMSVSGYTRBEIQDS 312

Qy 305 ISLWQYDHLTATYLLLLAKKA--RGKPVRL--LSSFCSCQASATP---QASATP---346  
Db 313 LVQQRNEVMATYLLGYKSSLEGGDTITLKPRPSADLTNSSAQFPSPSHKQVRSVSNPKQ 372

Qy 347 ---FTD-----IKSNWSLEDVTSADKNVAVGLIDYDWCEDDLSTGAATPRTSQTQKYWT 398  
Db 373 RRFSDQAGPAIPTSNYSYKKTOSNNAENKRP-----EEDRESGRKASSTAKVPA---S 422

Qy 399 ESNGVESKSLTP-----ALCRTPANKLNK---ENVYTPKSAVKN--EYFMPEPEPK---445  
Db 423 PUPGLERKKTTPSTNSVLSTSTNRSRNSPLLERASLGQASIQNGKDSLTPGSRAS 482

Qy 446 -----TPVKNQKHREILTT--PNRYTPPSKARNQC-----LKETPIKIPVNSTGDKLM 493  
Db 483 SASAASVAAARPRQKHQKMSASVHPKASGLPPTESNCEVPRESTAPQRPVSPSAHNIS 542

Qy 494 TGVISPER-----RCRSVELDLNQAHEETPK-----RKGA--KVF 527  
Db 543 SSGGAPDRTNPRGVSSRSTFHAGQLRQVR-DQQNLPGVTPASPSGHSQGRGASGSI 601



```
QY 471 RNQCLKE-----TPIKIPVNSTGTDKLMGTGVISSPER----- 501
Db 481 RNSPLDRLASLGQASIQNGKSDTAPQRPVSPSAHNISSSGAPDRTNFPRGVSSRSTF 540
QY 502 ---RCSRVELDLNQAHEETPK-----RKGA--KVFGS-----LERGL-----DKV 537
Db 541 HAQQLRQVR--DQONLPGVTVPASPGHSGOGRGASGSIKFTSKFVRRLNPEPKDRV 599
QY 538 ITVLTR-----SKRKGSGARDG--PRLKHLHYNTTTRLVNPOLLNEIMSILPKKHVDF 589
Db 600 ETLRPHVVGGGTDEKEEFREAKPRSLRFTWSMKTTSSMEPNEMREIRKVLDA NSCQS 659
QY 590 -VQGYTLKQOTSDFGKVTMQPELEVCOLEKQPDVVGIRQRLKGDWVYKRLVEDILSS 648
Db 660 ELHERYMLLCVHGTGPHENFVQWEMEVCVKLPRLSLNGVRFKRIISGTSMAFKNIASKIANE 719
QY 649 CKV 651
Db 720 LKL 722

RESULT 15
US-10-760-407-4
; Sequence 4, Application US/10760407
; Patent No. 6930173
; GENERAL INFORMATION:
; APPLICANT: YAN, Chunhua et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL001308-DIV II
; CURRENT APPLICATION NUMBER: US/10/760,407
; CURRENT FILING DATE: 2004-01-21
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 722
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-10-760-407-4
```

```
Query Match 22.0%; Score 759.5; DB 2; Length 722;
Best Local Similarity 29.7%; Pred. No. 3.5e-61;
Matches 215; Conservative 125; Mismatches 248; Indels 135; Gaps 23;

QY 11 YELHETIGTGPAKVKLACHILTGEMVAIKIMDKNTL-GSDLPRIKTEIEALKNLRHQHI 69
Db 53 YRLKLTIGKGNFAKVKLARHILITGKEVAVKIIDTQNLNSSLQKLPREVIMKVLNHPNI 112
QY 70 CQLYHLETANKIFMVLVEYCPGGELEFDYIIISODRLSEETRVVFRQIVSAVAVHSGGYA 129
Db 113 VKLFEVIEETKTLVLMVEYASGGEVFDYLVHGRMEKEARAKFRQIVSAVAVYCHQKFIV 172
QY 130 HRDLKPNILLFDEYHKLKLDIFGLCAK-PKGNKDYHLQTCGSLAYAAPELIQKSYLGS 188
Db 173 HRDLKAENLLDADNMKIDAFGFSNEFTGNK---LDTFCSGPPYAAAPELFQKKYDGP 229
QY 189 EADVWSMGILLYVMCGFLPPDDNNVVALYKIMRGKYDVPKWLSPSSILLLOQLQVDP 248
Db 230 EVDVWSGLVILYLVGSLPFDQGNKELRERVLRGKYRIPFYMTDCENLLKFLILNP 289
QY 249 KKRISMKNLNLHPWIMQDYNYPVEWOSKNPFIH----LDDDCVTELSVHRRNNRQTMEDL 304
Db 290 SKRGTLQEQIMKDRW---NVGHEDDELKPYVEPLPDYKDPRTTELWMSMGYTRBIIQDS 345
QY 305 ISLWQDHLTATYVLLLLAKA--RGKPVRL--RLSFSGQASATPFTDIKSNWSLEDV 360
Db 346 LVGQRYNVMAATYLLGLYKSSSELEGDTITLKPRLSADLTNSSAPSPSHKVQ-----RSV 399
QY 361 TASDKNVVAGLIDYDNCEDDLSTGAATPRTSQFTKYWTESNGVESK-----SLTP 410
Db 400 SANPKOR-----RSSDQVAPAIPTSNYSYK-TOSNNAENRRPEETGORKASSTA 448
```

Search completed: November 21, 2006, 14:15:15  
Job time: 56 secs



This Page Blank (uspto)

GenCore version 5.1.9  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: November 21, 2006, 14:14:30 ; Search time 188 Seconds  
(without alignments)  
1604.005 Million cell updates/sec

Title: US-10-656-598-2  
Perfect score: 3447  
Sequence: 1 MKDYDELLKYYELHETIGT.....KGDWVYKRLVEDILSSCKV 651

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 2097797 seqs, 463214858 residues

Total number of hits satisfying chosen parameters: 2097797

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

- Database : Published Applications AA Main:\*
- 1: /EMC\_Celerra\_SID33/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*
  - 2: /EMC\_Celerra\_SID33/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*
  - 3: /EMC\_Celerra\_SID33/ptodata/2/pubpaa/US09\_PUBCOMB.pep.\*
  - 4: /EMC\_Celerra\_SID33/ptodata/2/pubpaa/US10A\_PUBCOMB.pep.\*
  - 5: /EMC\_Celerra\_SID33/ptodata/2/pubpaa/US10B\_PUBCOMB.pep.\*
  - 6: /EMC\_Celerra\_SID33/ptodata/2/pubpaa/US11\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3447	100.0	651	3	US-09-870-937-10
2	3447	100.0	651	3	US-09-870-937-10
3	3447	100.0	651	3	US-09-870-937-10
4	3447	100.0	651	4	US-10-354-358-8
5	3447	100.0	651	4	US-10-295-027-506
6	3447	100.0	651	4	US-10-173-999-127
7	3447	100.0	651	5	US-10-656-598-2
8	3447	100.0	651	5	US-10-723-860-2444
9	3447	100.0	651	5	US-10-756-149-5359
10	3447	100.0	651	5	US-10-770-726-67
11	3447	100.0	651	6	US-11-177-138-10
12	3447	100.0	656	4	US-10-425-114-54360
13	3447	100.0	656	4	US-10-425-114-54261
14	1220	35.4	231	3	US-09-764-875-906
15	1176	34.1	272	4	US-10-106-698-5313
16	1156	33.5	726	4	US-10-369-493-5945
17	880	25.5	201	4	US-10-106-698-6277
18	791.5	23.0	713	5	US-10-995-561-881
19	791	22.9	744	3	US-09-835-081-4
20	789.5	22.9	667	4	US-10-016-248-74
21	789.5	22.9	713	4	US-10-161-565-27
22	789.5	22.9	729	5	US-10-995-561-878
23	788.5	22.9	729	4	US-10-142-356-11
24	788.5	22.9	729	4	US-10-195-101-33
25	788.5	22.9	729	4	US-10-161-565-26
26	788	22.9	744	3	US-09-319-585-3
27	788	22.9	744	5	US-10-995-561-876

28	787.5	22.8	737	5	US-10-995-561-880
29	785.5	22.8	733	5	US-10-995-561-877
30	772.5	22.4	731	5	US-10-491-467-27
31	770.5	22.4	691	3	US-09-919-585-6
32	770.5	22.4	724	3	US-09-919-585-9
33	769.5	22.3	691	5	US-10-756-149-5263
34	769.5	22.3	724	4	US-10-274-194-2
35	769.5	22.3	724	4	US-10-760-407-2
36	768	22.3	752	5	US-10-995-561-879
37	764.5	22.2	745	4	US-10-195-101-36
38	764.5	22.2	745	4	US-10-161-565-24
39	764.5	22.2	745	4	US-10-260-708-79
40	764.5	22.2	745	5	US-10-737-450-16
41	761	22.1	776	4	US-10-366-288-6
42	760	22.0	787	4	US-10-618-941-76
43	759.5	22.0	722	4	US-10-274-194-4
44	759.5	22.0	722	4	US-10-760-407-4
45	759.5	22.0	722	5	US-10-784-004-435

ALIGNMENTS

RESULT 1  
US-09-870-937-10  
; Sequence 10, Application US/09870937  
; Patent No. US20020049180A1  
; GENERAL INFORMATION:  
; APPLICANT: Wu, Bin  
; APPLICANT: Seeley, Todd  
; APPLICANT: Williams, Lewis T.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING NEOPLASTIC DISEASE USING  
; FILE REFERENCE: 200130.514/PP-01623.002  
; CURRENT APPLICATION NUMBER: US/09/870,937  
; CURRENT FILING DATE: 2001-05-30  
; NUMBER OF SEQ ID NOS: 12  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 10  
; TYPE: PRT  
; ORGANISM: Homo sapien  
US-09-870-937-10

Query Match	100.0%;	Score	3447;	DB	3;	Length	651;
Best Local Similarity	100.0%;	Pred. No.	7.8e-251;				
Matches	651;	Conservative	0;	Mismatches	0;	Indels	0;
QY	1	MKDYDELLKYYELHETIGTGGFAKVKLACHILTGEMVAIKINDKNTLGSDDLPRKTEIEA	60				
DB	1	MKDYDELLKYYELHETIGTGGFAKVKLACHILTGEMVAIKINDKNTLGSDDLPRKTEIEA	60				
QY	61	LKNLRHQCQLYHVLETKANKIFMWLEYPGCGELFDYIISQDRLSEETRVVPRQIVSAV	120				
DB	61	LKNLRHQCQLYHVLETKANKIFMWLEYPGCGELFDYIISQDRLSEETRVVPRQIVSAV	120				
QY	121	AYVHSGQYAHRLDKPENLLFDEYHKLKLDGFCAPKGNKDYHLQTCGSLAYAAPELI	180				
DB	121	AYVHSGQYAHRLDKPENLLFDEYHKLKLDGFCAPKGNKDYHLQTCGSLAYAAPELI	180				
QY	181	QKSYLGSEADYVWSMGILLIYVLMCGFLPDDDDNNVMAIYKIMRGKYDVPKWLSPSSILL	240				
DB	181	QKSYLGSEADYVWSMGILLIYVLMCGFLPDDDDNNVMAIYKIMRGKYDVPKWLSPSSILL	240				
QY	241	QOMLVDPKRIKSMKNLLNHPIMODYVPEWQSKNPFHLLDDCCVTELSVHHNNROT	300				
DB	241	QOMLVDPKRIKSMKNLLNHPIMODYVPEWQSKNPFHLLDDCCVTELSVHHNNROT	300				
QY	301	MEDLISLWQYDHLTATYLLLLAKAKGKPVRLRLSSFCGQASATPFTDIKNNWSLEDV	360				
DB	301	MEDLISLWQYDHLTATYLLLLAKAKGKPVRLRLSSFCGQASATPFTDIKNNWSLEDV	360				
QY	361	TASDKNYVAGLIDYDWCEDDLSTGAATPRTSFTKYTWESNGVESKSLTPALCRTPANKL	420				

Db 361 TASDKNYVAGLIDYDWCEDDLSTGAATPRTSQFTKYWTESNGVESKSLTPALCRTPANKL 420  
QY 421 KKNENYVTPKSAVKNEEYFMFPPEKTPVNKNQHKREILTPNRYTTPSKARNOCLKETPI 480  
Db 421 KKNENYVTPKSAVKNEEYFMFPPEKTPVNKNQHKREILTPNRYTTPSKARNOCLKETPI 480  
QY 481 KIPVNSTGTDKMTGTVISPERCRSVELDLNQAAMEETPKRGAKVFGSLERGLDKVITV 540  
Db 481 KIPVNSTGTDKMTGTVISPERCRSVELDLNQAAMEETPKRGAKVFGSLERGLDKVITV 540  
QY 541 LTRSKRGSGARDGPRRLKHLHYNTTTLVNPDPQLNEIMSILPKKHVDVFKGYTLKCO 600  
Db 541 LTRSKRGSGARDGPRRLKHLHYNTTTLVNPDPQLNEIMSILPKKHVDVFKGYTLKCO 600  
QY 601 QSDFGKVTWQFLEVCQLOKPDVVGIRRQRLKGDAMVYKRLVEDILSSCKV 651  
Db 601 QSDFGKVTWQFLEVCQLOKPDVVGIRRQRLKGDAMVYKRLVEDILSSCKV 651

RESULT 2  
US-09-974-298-112  
; Sequence 112, Application US/09974298  
; Patent No. US20020156263A1  
; GENERAL INFORMATION:  
; APPLICANT: Chen, Hwei-Mei  
; TITLE OF INVENTION: GENES EXPRESSED IN BREAST CANCER  
; FILE REFERENCE: PA-0037 P  
; CURRENT APPLICATION NUMBER: US/09/974,298  
; PRIOR FILING DATE: 2001-10-04  
; PRIOR APPLICATION NUMBER: 60/238,331  
; PRIOR FILING DATE: 2000-05-10  
; NUMBER OF SEQ ID NOS: 194  
; SOFTWARE: PERL Program  
; SEQ ID NO 112  
; TYPE: PRT  
; LENGTH: 651  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc\_feature  
; OTHER INFORMATION: Incyte ID No. US20020156263A1 256716CD1  
US-09-974-298-112

Query Match 100.0%; Score 3447; DB 3; Length 651;  
Best Local Similarity 100.0%; Pred. No. 7.8e-251;  
Matches 651; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKDYDELLKYYELHETIGTGGFAKVKLACHILTGEMVAIKIMDKNTLGSDLPRIKTEIEA 60  
Db 1 MKDYDELLKYYELHETIGTGGFAKVKLACHILTGEMVAIKIMDKNTLGSDLPRIKTEIEA 60  
QY 61 LKNLRHQHICQLYHVLETANKIPWVLEYCPGGELFDYIISQDRLSBEETRVVFRQIVSAV 120  
Db 61 LKNLRHQHICQLYHVLETANKIPWVLEYCPGGELFDYIISQDRLSBEETRVVFRQIVSAV 120  
QY 121 AVVHSQGYAHRDLKPENLLFDEYHKLIDFGICAKPKGNKDVLQTCGSLAYAAPELI 180  
Db 121 AVVHSQGYAHRDLKPENLLFDEYHKLIDFGICAKPKGNKDVLQTCGSLAYAAPELI 180  
QY 181 QGKSILGSEADVMSGILLVYLMCGFLPFDDNNMALYKKIMRGKYDVPKWLSPSSILL 240  
Db 181 QGKSILGSEADVMSGILLVYLMCGFLPFDDNNMALYKKIMRGKYDVPKWLSPSSILL 240  
QY 241 QOMLOVDPKKRISMKNLNHPWIMQDYNYPVEWQSKNPPFHLDDDCVTEL SVHHRNRQT 300  
Db 241 QOMLOVDPKKRISMKNLNHPWIMQDYNYPVEWQSKNPPFHLDDDCVTEL SVHHRNRQT 300  
QY 301 MEDLISLWQYDHLTATYVLLLLAKARGKPVRLRLSSFCGQASATPTDIIKSNWSLEDV 360  
Db 301 MEDLISLWQYDHLTATYVLLLLAKARGKPVRLRLSSFCGQASATPTDIIKSNWSLEDV 360  
QY 361 TASDKNYVAGLIDYDWCEDDLSTGAATPRTSQFTKYWTESNGVESKSLTPALCRTPANKL 420

Db 361 TASDKNYVAGLIDYDWCEDDLSTGAATPRTSQFTKYWTESNGVESKSLTPALCRTPANKL 420  
QY 421 KKNENYVTPKSAVKNEEYFMFPPEKTPVNKNQHKREILTPNRYTTPSKARNOCLKETPI 480  
Db 421 KKNENYVTPKSAVKNEEYFMFPPEKTPVNKNQHKREILTPNRYTTPSKARNOCLKETPI 480  
QY 481 KIPVNSTGTDKMTGTVISPERCRSVELDLNQAAMEETPKRGAKVFGSLERGLDKVITV 540  
Db 481 KIPVNSTGTDKMTGTVISPERCRSVELDLNQAAMEETPKRGAKVFGSLERGLDKVITV 540  
QY 541 LTRSKRGSGARDGPRRLKHLHYNTTTLVNPDPQLNEIMSILPKKHVDVFKGYTLKCO 600  
Db 541 LTRSKRGSGARDGPRRLKHLHYNTTTLVNPDPQLNEIMSILPKKHVDVFKGYTLKCO 600  
QY 601 QSDFGKVTWQFLEVCQLOKPDVVGIRRQRLKGDAMVYKRLVEDILSSCKV 651  
Db 601 QSDFGKVTWQFLEVCQLOKPDVVGIRRQRLKGDAMVYKRLVEDILSSCKV 651

RESULT 3  
US-10-354-358-8  
; Sequence 8, Application US/10354358  
; Publication No. US20030157082A1  
; GENERAL INFORMATION:  
; APPLICANT: Millennium Pharmaceuticals, Inc  
; APPLICANT: Hunter, John Joseph  
; APPLICANT: MacBeth, Kyle J.  
; APPLICANT: Tsai, Fong-Ying  
; APPLICANT: Lesoon, Andrea  
; APPLICANT: Lightcap, Eric S.  
; APPLICANT: Williamson, Mark  
; APPLICANT: Rudolph-Owen, Laura A.  
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING  
; TITLE OF INVENTION: CANCER USING 140, 1470, 1686, 2089, 2427, 3702, 5891, 6428,  
; TITLE OF INVENTION: 7181, 7650, 25641, 69583, 49863, 8897, 1682, 17667, 9235,  
; TITLE OF INVENTION: 3703, 14171, 10359, 1660, 1450, 18894, 2088, 32427, 2160,  
; TITLE OF INVENTION: 9252, 9389, 1642, 85269, 10297, 1584, 9525, 14124, 4469,  
; TITLE OF INVENTION: 8990, 2100, 9288, 64698, 10480, 20893, 33230, 1586, 9943,  
; TITLE OF INVENTION: 16334, 68862, 9011, 14031, 6178, 21225, 1420, 32236, 2099,  
; TITLE OF INVENTION: 2150, 26583, 2784, 8941, 9811, 27444, 50566 OR 66428 MOLECULES  
; FILE REFERENCE: MP102-020PIRNONIM  
; CURRENT APPLICATION NUMBER: US/10/354,358  
; CURRENT FILING DATE: 2003-01-30  
; PRIOR APPLICATION NUMBER: US 60/353,600  
; PRIOR FILING DATE: 2002-01-31  
; PRIOR APPLICATION NUMBER: US 60/364,517  
; PRIOR FILING DATE: 2002-03-15  
; PRIOR APPLICATION NUMBER: US 60/371,075  
; PRIOR FILING DATE: 2002-04-09  
; PRIOR APPLICATION NUMBER: US 60/371,507  
; PRIOR FILING DATE: 2002-04-10  
; PRIOR APPLICATION NUMBER: US 60/372,984  
; PRIOR FILING DATE: 2002-04-16  
; PRIOR APPLICATION NUMBER: US 60/374,194  
; PRIOR FILING DATE: 2002-04-19  
; PRIOR APPLICATION NUMBER: US 60/382,995  
; PRIOR FILING DATE: 2002-05-24  
; PRIOR APPLICATION NUMBER: US 60/385,023  
; PRIOR FILING DATE: 2002-05-31  
; PRIOR APPLICATION NUMBER: US 60/388,853  
; PRIOR FILING DATE: 2002-06-14  
; PRIOR APPLICATION NUMBER: US 60/389,395  
; PRIOR FILING DATE: 2002-06-17  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 122  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 8  
; LENGTH: 651  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-354-358-8

Query Match 100.0%; Score 3447; DB 4; Length 651;

Best Local Similarity 100.0%; Pred. No. 7.8e-251;  
Matches 651; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKDYDELLKYELHETIGTGFAKVLACHILTGEMVAIKIMDKNTLGSDDLPRIKTEIEA 60  
DB 1 MKDYDELLKYELHETIGTGFAKVLACHILTGEMVAIKIMDKNTLGSDDLPRIKTEIEA 60  
QY 61 LKNLRHQHICOLYHVLETTANKIFMWLEVCYCGGELFDYIISQDRLSEETRVVFRQIVSAV 120  
DB 61 LKNLRHQHICOLYHVLETTANKIFMWLEVCYCGGELFDYIISQDRLSEETRVVFRQIVSAV 120  
QY 121 AYVHSGQYAHARDLKPENLLFDEYHKLKIDFGLCAKPKGNKDYHLQTCGSLAYAAPELI 180  
DB 121 AYVHSGQYAHARDLKPENLLFDEYHKLKIDFGLCAKPKGNKDYHLQTCGSLAYAAPELI 180  
QY 181 QGKSYLGSEADVMSGILLYVLMCGFLPDDDDNMALYKIMRGKYDVPKWLSPSSILL 240  
DB 181 QGKSYLGSEADVMSGILLYVLMCGFLPDDDDNMALYKIMRGKYDVPKWLSPSSILL 240  
QY 241 QOMLQVDPKCRISMKNLNHPIMQDYNYPVEWQSKNPFILHDDDCVTELSVHHRNROT 300  
DB 241 QOMLQVDPKCRISMKNLNHPIMQDYNYPVEWQSKNPFILHDDDCVTELSVHHRNROT 300  
QY 301 MEDLISLWQYDHLTATYLLALLAKARGKPVRLRLSSFCGQASATPPTDIKNNWSLEDV 360  
DB 301 MEDLISLWQYDHLTATYLLALLAKARGKPVRLRLSSFCGQASATPPTDIKNNWSLEDV 360  
QY 361 TASDKNYVAGLIDYDWCEDDLSTGAATPRTSQTKYWTESNGVESKSLTPALCRTPANKL 420  
DB 361 TASDKNYVAGLIDYDWCEDDLSTGAATPRTSQTKYWTESNGVESKSLTPALCRTPANKL 420  
QY 421 KKNENYVTPKSAVKNBEYFMPEPKTPVNKQHKREILITPNRYTTPSKARNQCLKETPI 480  
DB 421 KKNENYVTPKSAVKNBEYFMPEPKTPVNKQHKREILITPNRYTTPSKARNQCLKETPI 480  
QY 481 KIPVNSTGDKLMTGVI SPERRCRSVELDLNQAHEETPKRGAKVFGSLERGLDKVITV 540  
DB 481 KIPVNSTGDKLMTGVI SPERRCRSVELDLNQAHEETPKRGAKVFGSLERGLDKVITV 540  
QY 541 LTRSRKKSARDGPRRLKLYHNVTTTLLVNPDLNNEIMSLPKKHVDVQKGYTLKCOY 600  
DB 541 LTRSRKKSARDGPRRLKLYHNVTTTLLVNPDLNNEIMSLPKKHVDVQKGYTLKCOY 600  
QY 601 QSDFGKVTMQFELEVQOLQKPDVVGIRRQRLKGDAMVYKRLVEDILSSCKV 651  
DB 601 QSDFGKVTMQFELEVQOLQKPDVVGIRRQRLKGDAMVYKRLVEDILSSCKV 651

RESULT 4  
US-10-295-027-506  
; Sequence 506, Application US/10295027  
; Publication No. US20030232350A1  
; GENERAL INFORMATION:  
; APPLICANT: Afar, Daniel  
; APPLICANT: Aziz, Nataasha  
; APPLICANT: Ginsberg, Wendy M.  
; APPLICANT: Gish, Kurt C.  
; APPLICANT: Glynn, Richard  
; APPLICANT: Heverzi, Peter A.  
; APPLICANT: Mack, David H.  
; APPLICANT: Murray, Richard  
; APPLICANT: Watson, Susan R.  
; APPLICANT: Eos Biotechnology, Inc.  
; TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and  
; FILE REFERENCE: 018501-012500S  
; CURRENT FILING DATE: 2002-11-13  
; CURRENT APPLICATION NUMBER: US/09/295,027  
; PRIOR FILING DATE: 2000-09-15  
; PRIOR APPLICATION NUMBER: US 09/663,733  
; PRIOR FILING DATE: 2000-09-15  
; PRIOR APPLICATION NUMBER: US 60/350,666  
; PRIOR FILING DATE: 2001-11-13  
; PRIOR APPLICATION NUMBER: US 60/335,394

Query Match 100.0%; Score 3447; DB 4; Length 651;  
Best Local Similarity 100.0%; Pred. No. 7.8e-251;  
Matches 651; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKDYDELLKYELHETIGTGFAKVLACHILTGEMVAIKIMDKNTLGSDDLPRIKTEIEA 60  
DB 1 MKDYDELLKYELHETIGTGFAKVLACHILTGEMVAIKIMDKNTLGSDDLPRIKTEIEA 60  
QY 61 LKNLRHQHICOLYHVLETTANKIFMWLEVCYCGGELFDYIISQDRLSEETRVVFRQIVSAV 120  
DB 61 LKNLRHQHICOLYHVLETTANKIFMWLEVCYCGGELFDYIISQDRLSEETRVVFRQIVSAV 120  
QY 121 AYVHSGQYAHARDLKPENLLFDEYHKLKIDFGLCAKPKGNKDYHLQTCGSLAYAAPELI 180  
DB 121 AYVHSGQYAHARDLKPENLLFDEYHKLKIDFGLCAKPKGNKDYHLQTCGSLAYAAPELI 180  
QY 181 QGKSYLGSEADVMSGILLYVLMCGFLPDDDDNMALYKIMRGKYDVPKWLSPSSILL 240  
DB 181 QGKSYLGSEADVMSGILLYVLMCGFLPDDDDNMALYKIMRGKYDVPKWLSPSSILL 240  
QY 241 QOMLQVDPKCRISMKNLNHPIMQDYNYPVEWQSKNPFILHDDDCVTELSVHHRNROT 300  
DB 241 QOMLQVDPKCRISMKNLNHPIMQDYNYPVEWQSKNPFILHDDDCVTELSVHHRNROT 300  
QY 301 MEDLISLWQYDHLTATYLLALLAKARGKPVRLRLSSFCGQASATPPTDIKNNWSLEDV 360  
DB 301 MEDLISLWQYDHLTATYLLALLAKARGKPVRLRLSSFCGQASATPPTDIKNNWSLEDV 360  
QY 361 TASDKNYVAGLIDYDWCEDDLSTGAATPRTSQTKYWTESNGVESKSLTPALCRTPANKL 420  
DB 361 TASDKNYVAGLIDYDWCEDDLSTGAATPRTSQTKYWTESNGVESKSLTPALCRTPANKL 420  
QY 421 KKNENYVTPKSAVKNBEYFMPEPKTPVNKQHKREILITPNRYTTPSKARNQCLKETPI 480  
DB 421 KKNENYVTPKSAVKNBEYFMPEPKTPVNKQHKREILITPNRYTTPSKARNQCLKETPI 480  
QY 481 KIPVNSTGDKLMTGVI SPERRCRSVELDLNQAHEETPKRGAKVFGSLERGLDKVITV 540  
DB 481 KIPVNSTGDKLMTGVI SPERRCRSVELDLNQAHEETPKRGAKVFGSLERGLDKVITV 540  
QY 541 LTRSRKKSARDGPRRLKLYHNVTTTLLVNPDLNNEIMSLPKKHVDVQKGYTLKCOY 600  
DB 541 LTRSRKKSARDGPRRLKLYHNVTTTLLVNPDLNNEIMSLPKKHVDVQKGYTLKCOY 600  
QY 601 QSDFGKVTMQFELEVQOLQKPDVVGIRRQRLKGDAMVYKRLVEDILSSCKV 651  
DB 601 QSDFGKVTMQFELEVQOLQKPDVVGIRRQRLKGDAMVYKRLVEDILSSCKV 651

RESULT 5  
US-10-173-999-127  
; Sequence 127, Application US/10173999  
; Publication No. US20040005563A1  
; GENERAL INFORMATION:  
; APPLICANT: Mack, David H.  
; APPLICANT: Gish, Kurt C.  
; APPLICANT: Eos Biotechnology, Inc.  
; TITLE OF INVENTION: Methods of Diagnosis of Ovarian Cancer, Compositions  
; TITLE OF INVENTION: and Methods of Screening for Modulators of Ovarian  
; TITLE OF INVENTION: Cancer  
; FILE REFERENCE: 018501-002420US  
; CURRENT APPLICATION NUMBER: US/10/173,999  
; CURRENT FILING DATE: 2002-06-17  
; PRIOR APPLICATION NUMBER: US 60/299,234  
; PRIOR FILING DATE: 2001-06-18  
; PRIOR APPLICATION NUMBER: US 60/315,287  
; PRIOR FILING DATE: 2001-08-27  
; PRIOR APPLICATION NUMBER: US 60/350,666  
; PRIOR FILING DATE: 2001-11-13  
; PRIOR APPLICATION NUMBER: US 60/372,246  
; PRIOR FILING DATE: 2001-04-12  
; NUMBER OF SEQ ID NOS: 163  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 127  
; LENGTH: 651  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-173-999-127

Query Match 100.0%; Score 3447; DB 4; Length 651;  
Best Local Similarity 100.0%; Pred. No. 7.8e-251;  
Matches 651; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MKDYDELLKYYELHETIGTGGAFAKVLACHILTGEMVAIKIMDKNTLGSGLDPRIKTEIA 60  
Db 1 MKDYDELLKYYELHETIGTGGAFAKVLACHILTGEMVAIKIMDKNTLGSGLDPRIKTEIA 60

Qy 61 LKNLRHQHICOLYHVLLETANKIFMWLEPCPGGELFDYIIISODRLSEETRVVFRQIVSAV 120  
Db 61 LKNLRHQHICOLYHVLLETANKIFMWLEPCPGGELFDYIIISODRLSEETRVVFRQIVSAV 120

Qy 121 AYVHSGQYAHARDLKPENLLFDEYHKLKIDFGLCAKPGKNDYHLOTCCGSLAYAAPELI 180  
Db 121 AYVHSGQYAHARDLKPENLLFDEYHKLKIDFGLCAKPGKNDYHLOTCCGSLAYAAPELI 180

Qy 181 QGKSYLGSEADVMSGILLYVLMCGFLPFDDNNMVALYKKIMRGKYDVPKWLSPSSILL 240  
Db 181 QGKSYLGSEADVMSGILLYVLMCGFLPFDDNNMVALYKKIMRGKYDVPKWLSPSSILL 240

Qy 241 QQMLQVDPKKRISMKNLLNHPIMQDYNYPVWQSKNPFILHDDDCVTLSVHHRNROT 300  
Db 241 QQMLQVDPKKRISMKNLLNHPIMQDYNYPVWQSKNPFILHDDDCVTLSVHHRNROT 300

Qy 301 MEDLISLWQYDHLTATYLLLLAKKARGKPVRLRLSSFCGQASATPFTDIKSNWNSLEDV 360  
Db 301 MEDLISLWQYDHLTATYLLLLAKKARGKPVRLRLSSFCGQASATPFTDIKSNWNSLEDV 360

Qy 361 TASDKNYVAGLIDYDWCEDDLSTGAATPRTSQFTKYWTESNGVESKSLTPALCRTPANKL 420  
Db 361 TASDKNYVAGLIDYDWCEDDLSTGAATPRTSQFTKYWTESNGVESKSLTPALCRTPANKL 420

Qy 421 KKNENYVTPKSAVKNBEYFMFPPEKTPVNKNQHKREILITPNRYTTPSKARNQCLKETPI 480  
Db 421 KKNENYVTPKSAVKNBEYFMFPPEKTPVNKNQHKREILITPNRYTTPSKARNQCLKETPI 480

Qy 481 KIPVNSTGTDKLMGTGVI SPERRCRSVELDLNQAHEETPKRGAKVFGSLERGLDKVITV 540  
Db 481 KIPVNSTGTDKLMGTGVI SPERRCRSVELDLNQAHEETPKRGAKVFGSLERGLDKVITV 540

Qy 541 LTRSKRKSARGDGRRLKLYHNVTTTTLVNPDLQLLNEIMSLPKGHVDFVQKGYTLKCQT 600  
Db 541 LTRSKRKSARGDGRRLKLYHNVTTTTLVNPDLQLLNEIMSLPKGHVDFVQKGYTLKCQT 600

Qy 601 QSDFGKVTMQELEYCQLQKPDVVGIRQRQLKGDAAWYKRLVEDILSSCKV 651  
Db 601 QSDFGKVTMQELEYCQLQKPDVVGIRQRQLKGDAAWYKRLVEDILSSCKV 651

RESULT 6  
US-10-656-598-2  
; Sequence 2, Application US/10656598  
; Publication No. US20040229232A1  
; GENERAL INFORMATION:  
; APPLICANT: DAVID P. DAVIS  
; APPLICANT: FREDERIC J. DESAUVAGE  
; APPLICANT: WILLIAM I. WOOD  
; APPLICANT: ZEMIN ZHANG  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DIAGNOSIS AND  
; TITLE OF INVENTION: TREATMENT OF TUMOR  
; FILE REFERENCE: P1981R1P1-US  
; CURRENT APPLICATION NUMBER: US/10/656,598  
; CURRENT FILING DATE: 2003-09-05  
; PRIOR APPLICATION NUMBER: US 60/410,166  
; PRIOR FILING DATE: 2002-09-11  
; NUMBER OF SEQ ID NOS: 52  
; SEQ ID NO 2  
; LENGTH: 651  
; TYPE: PRT  
; ORGANISM: Homo sapien  
US-10-656-598-2

Query Match 100.0%; Score 3447; DB 5; Length 651;  
Best Local Similarity 100.0%; Pred. No. 7.8e-251;  
Matches 651; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MKDYDELLKYYELHETIGTGGAFAKVLACHILTGEMVAIKIMDKNTLGSGLDPRIKTEIA 60  
Db 1 MKDYDELLKYYELHETIGTGGAFAKVLACHILTGEMVAIKIMDKNTLGSGLDPRIKTEIA 60

Qy 61 LKNLRHQHICOLYHVLLETANKIFMWLEPCPGGELFDYIIISODRLSEETRVVFRQIVSAV 120  
Db 61 LKNLRHQHICOLYHVLLETANKIFMWLEPCPGGELFDYIIISODRLSEETRVVFRQIVSAV 120

Qy 121 AYVHSGQYAHARDLKPENLLFDEYHKLKIDFGLCAKPGKNDYHLOTCCGSLAYAAPELI 180  
Db 121 AYVHSGQYAHARDLKPENLLFDEYHKLKIDFGLCAKPGKNDYHLOTCCGSLAYAAPELI 180

Qy 181 QGKSYLGSEADVMSGILLYVLMCGFLPFDDNNMVALYKKIMRGKYDVPKWLSPSSILL 240  
Db 181 QGKSYLGSEADVMSGILLYVLMCGFLPFDDNNMVALYKKIMRGKYDVPKWLSPSSILL 240

Qy 241 QQMLQVDPKKRISMKNLLNHPIMQDYNYPVWQSKNPFILHDDDCVTLSVHHRNROT 300  
Db 241 QQMLQVDPKKRISMKNLLNHPIMQDYNYPVWQSKNPFILHDDDCVTLSVHHRNROT 300

Qy 301 MEDLISLWQYDHLTATYLLLLAKKARGKPVRLRLSSFCGQASATPFTDIKSNWNSLEDV 360  
Db 301 MEDLISLWQYDHLTATYLLLLAKKARGKPVRLRLSSFCGQASATPFTDIKSNWNSLEDV 360

Qy 361 TASDKNYVAGLIDYDWCEDDLSTGAATPRTSQFTKYWTESNGVESKSLTPALCRTPANKL 420  
Db 361 TASDKNYVAGLIDYDWCEDDLSTGAATPRTSQFTKYWTESNGVESKSLTPALCRTPANKL 420

Qy 421 KKNENYVTPKSAVKNBEYFMFPPEKTPVNKNQHKREILITPNRYTTPSKARNQCLKETPI 480  
Db 421 KKNENYVTPKSAVKNBEYFMFPPEKTPVNKNQHKREILITPNRYTTPSKARNQCLKETPI 480

Qy 481 KIPVNSTGTDKLMGTGVI SPERRCRSVELDLNQAHEETPKRGAKVFGSLERGLDKVITV 540  
Db 481 KIPVNSTGTDKLMGTGVI SPERRCRSVELDLNQAHEETPKRGAKVFGSLERGLDKVITV 540

Qy 541 LTRSKRKSARGDGRRLKLYHNVTTTTLVNPDLQLLNEIMSLPKGHVDFVQKGYTLKCQT 600  
Db 541 LTRSKRKSARGDGRRLKLYHNVTTTTLVNPDLQLLNEIMSLPKGHVDFVQKGYTLKCQT 600

QY 601 QSDFGKVTMQFELEVCOQKPDVVGIRRQRLKGDWVYKRLVEDILSSCKV 651  
Db 601 QSDFGKVTMQFELEVCOQKPDVVGIRRQRLKGDWVYKRLVEDILSSCKV 651

RESULT 7  
US-10-723-860-2444  
; Sequence 2444, Application US/10723860  
; Publication No. US20040253606A1  
; GENERAL INFORMATION:  
; APPLICANT: Aziz, Natasha  
; APPLICANT: Gineburg, Wendy M.  
; APPLICANT: Zlotnik, Albert  
; TITLE OF INVENTION: Methods of Diagnosis of Soft Tissue Sarcoma, Compositions &  
; FILE REFERENCE: 05892.0193.NPUS01  
; CURRENT APPLICATION NUMBER: US/10/723,860  
; CURRENT FILING DATE: 2003-11-26  
; PRIOR APPLICATION NUMBER: 60/429,739  
; PRIOR FILING DATE: 2002-11-26  
; NUMBER OF SEQ ID NOS: 8393  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 2444  
; LENGTH: 651  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-723-860-2444

Query Match 100.0%; Score 3447; DB 5; Length 651;  
Best Local Similarity 100.0%; Pred. No. 7.8e-251;  
Matches 651; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKDYDELLKYYELHETIGTGGAFAKVLACHILTGEMVAIKMDKNTLGSGLDLPRIKTEIEA 60  
Db 1 MKDYDELLKYYELHETIGTGGAFAKVLACHILTGEMVAIKMDKNTLGSGLDLPRIKTEIEA 60  
QY 61 LKNLRHQHICQLYHVLETANKIFMWLEYCPGGELFDYIISQRLSEETRVVFRQIVSAV 120  
Db 61 LKNLRHQHICQLYHVLETANKIFMWLEYCPGGELFDYIISQRLSEETRVVFRQIVSAV 120  
QY 121 AYVHSGYGAHRDLKPNELLFDEYHKLLIDFGLCAKPKGNKDYHLQTCGSLAYAAPELI 180  
Db 121 AYVHSGYGAHRDLKPNELLFDEYHKLLIDFGLCAKPKGNKDYHLQTCGSLAYAAPELI 180  
QY 181 QGKSYLGSSEADVWMSGILLVYLMCGFLPFDDDDNNVALYKKIMRGKYDVPKWLSPSSILL 240  
Db 181 QGKSYLGSSEADVWMSGILLVYLMCGFLPFDDDDNNVALYKKIMRGKYDVPKWLSPSSILL 240  
QY 241 QQMLQVDPKKRISMKNLLNHPWIMQDYNYPVWQSKNPFHLLDDDCVTLSVHHRNROT 300  
Db 241 QQMLQVDPKKRISMKNLLNHPWIMQDYNYPVWQSKNPFHLLDDDCVTLSVHHRNROT 300  
QY 301 MEDLISLWQYDHLTATYLLALLAKKARGKPVRLRLSSFCGQASATPFTDIKSNWNSLEDV 360  
Db 301 MEDLISLWQYDHLTATYLLALLAKKARGKPVRLRLSSFCGQASATPFTDIKSNWNSLEDV 360  
QY 361 TASDKNYVAGLIDYDWCEDDLSGTAAATPRTSQFTKYWTESNGVESKSLTPALCRTPANKL 420  
Db 361 TASDKNYVAGLIDYDWCEDDLSGTAAATPRTSQFTKYWTESNGVESKSLTPALCRTPANKL 420  
QY 421 KKNENYVTPKSAVKNEEYFMPPEKTPVNKNQHKREILTPNRYTTPSKARNOCLKETPI 480  
Db 421 KKNENYVTPKSAVKNEEYFMPPEKTPVNKNQHKREILTPNRYTTPSKARNOCLKETPI 480  
QY 481 KIPVNSTGDKLMTGVIISPERRCRSVELDLNOAHMEETPKRGAKVFGSLERGLDKVITV 540  
Db 481 KIPVNSTGDKLMTGVIISPERRCRSVELDLNOAHMEETPKRGAKVFGSLERGLDKVITV 540  
QY 541 LTRSRKKSARGDPRLKHLHYNTTTRLVNPDLQLINEIMSLPKKHVDVFOQGYTLKCO 600  
Db 541 LTRSRKKSARGDPRLKHLHYNTTTRLVNPDLQLINEIMSLPKKHVDVFOQGYTLKCO 600  
QY 601 QSDFGKVTMQFELEVCOQKPDVVGIRRQRLKGDWVYKRLVEDILSSCKV 651

Db 601 QSDFGKVTMQFELEVCOQKPDVVGIRRQRLKGDWVYKRLVEDILSSCKV 651

RESULT 8  
US-10-756-149-5359  
; Sequence 5359, Application US/10756149  
; Publication No. US20050181375A1  
; GENERAL INFORMATION:  
; APPLICANT: Aziz, Natasha  
; APPLICANT: Zlotnik, Albert  
; TITLE OF INVENTION: NOVEL METHODS OF DIAGNOSIS OF METASTATIC CANCER, COMPOSITIONS AND  
; FILE REFERENCE: file  
; CURRENT APPLICATION NUMBER: US/10/756,149  
; CURRENT FILING DATE: 2004-01-12  
; NUMBER OF SEQ ID NOS: 5818  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 5359  
; LENGTH: 651  
; TYPE: PRT  
; ORGANISM: Homo Sapiens  
US-10-756-149-5359

Query Match 100.0%; Score 3447; DB 5; Length 651;  
Best Local Similarity 100.0%; Pred. No. 7.8e-251;  
Matches 651; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKDYDELLKYYELHETIGTGGAFAKVLACHILTGEMVAIKMDKNTLGSGLDLPRIKTEIEA 60  
Db 1 MKDYDELLKYYELHETIGTGGAFAKVLACHILTGEMVAIKMDKNTLGSGLDLPRIKTEIEA 60  
QY 61 LKNLRHQHICQLYHVLETANKIFMWLEYCPGGELFDYIISQRLSEETRVVFRQIVSAV 120  
Db 61 LKNLRHQHICQLYHVLETANKIFMWLEYCPGGELFDYIISQRLSEETRVVFRQIVSAV 120  
QY 121 AYVHSGYGAHRDLKPNELLFDEYHKLLIDFGLCAKPKGNKDYHLQTCGSLAYAAPELI 180  
Db 121 AYVHSGYGAHRDLKPNELLFDEYHKLLIDFGLCAKPKGNKDYHLQTCGSLAYAAPELI 180  
QY 181 QGKSYLGSSEADVWMSGILLVYLMCGFLPFDDDDNNVALYKKIMRGKYDVPKWLSPSSILL 240  
Db 181 QGKSYLGSSEADVWMSGILLVYLMCGFLPFDDDDNNVALYKKIMRGKYDVPKWLSPSSILL 240  
QY 241 QQMLQVDPKKRISMKNLLNHPWIMQDYNYPVWQSKNPFHLLDDDCVTLSVHHRNROT 300  
Db 241 QQMLQVDPKKRISMKNLLNHPWIMQDYNYPVWQSKNPFHLLDDDCVTLSVHHRNROT 300  
QY 301 MEDLISLWQYDHLTATYLLALLAKKARGKPVRLRLSSFCGQASATPFTDIKSNWNSLEDV 360  
Db 301 MEDLISLWQYDHLTATYLLALLAKKARGKPVRLRLSSFCGQASATPFTDIKSNWNSLEDV 360  
QY 361 TASDKNYVAGLIDYDWCEDDLSGTAAATPRTSQFTKYWTESNGVESKSLTPALCRTPANKL 420  
Db 361 TASDKNYVAGLIDYDWCEDDLSGTAAATPRTSQFTKYWTESNGVESKSLTPALCRTPANKL 420  
QY 421 KKNENYVTPKSAVKNEEYFMPPEKTPVNKNQHKREILTPNRYTTPSKARNOCLKETPI 480  
Db 421 KKNENYVTPKSAVKNEEYFMPPEKTPVNKNQHKREILTPNRYTTPSKARNOCLKETPI 480  
QY 481 KIPVNSTGDKLMTGVIISPERRCRSVELDLNOAHMEETPKRGAKVFGSLERGLDKVITV 540  
Db 481 KIPVNSTGDKLMTGVIISPERRCRSVELDLNOAHMEETPKRGAKVFGSLERGLDKVITV 540  
QY 541 LTRSRKKSARGDPRLKHLHYNTTTRLVNPDLQLINEIMSLPKKHVDVFOQGYTLKCO 600  
Db 541 LTRSRKKSARGDPRLKHLHYNTTTRLVNPDLQLINEIMSLPKKHVDVFOQGYTLKCO 600  
QY 601 QSDFGKVTMQFELEVCOQKPDVVGIRRQRLKGDWVYKRLVEDILSSCKV 651  
Db 601 QSDFGKVTMQFELEVCOQKPDVVGIRRQRLKGDWVYKRLVEDILSSCKV 651

```
RESULT 9
US-10-770-726-67
; Sequence 67, Application US/10770726
; Publication No. US20050266409A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Brown, Eugene
; APPLICANT: Liu, Wei
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR DIAGNOSING, PREVENTING, AND TREATING
; FILE REFERENCE: AM101079 (031896-010000)
; CURRENT APPLICATION NUMBER: US/10/770,726
; CURRENT FILING DATE: 2004-02-04
; NUMBER OF SEQ ID NOS: 48640
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 67
; LENGTH: 651
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-770-726-67

Query Match      100.0%; Score 3447; DB 5; Length 651;
Best Local Similarity 100.0%; Pred. No. 7.8e-251;
Matches 651; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKDYDELLKYYELHETIGTGPAKVKLACHILTGEMVAIKIMDKNTLGSDDLPRIKTEIEA 60
DB 1 MKDYDELLKYYELHETIGTGPAKVKLACHILTGEMVAIKIMDKNTLGSDDLPRIKTEIEA 60
QY 61 LKNLRHQHICQLYHVLETANKIFMWLEYCPGGELFDYIISQDRLSEETRVVFRQIVSAV 120
DB 61 LKNLRHQHICQLYHVLETANKIFMWLEYCPGGELFDYIISQDRLSEETRVVFRQIVSAV 120
QY 121 AYVHSQGYAHRDLKPNLLFDEYHKLIDFGLCAKPGKNDYHLQTCGSLAYAAPELI 180
DB 121 AYVHSQGYAHRDLKPNLLFDEYHKLIDFGLCAKPGKNDYHLQTCGSLAYAAPELI 180
QY 181 QGKSYLGSEADVWSMGILLVLMCGFLPDDDDNMALYKKIMRGKYDVPKWLSPSSILL 240
DB 181 QGKSYLGSEADVWSMGILLVLMCGFLPDDDDNMALYKKIMRGKYDVPKWLSPSSILL 240
QY 241 QQMLQVDPKKRISMKNLLNHPWIMQDYNYPVEWQSKNPFILHDDDCVTLSVHHRNNROT 300
DB 241 QQMLQVDPKKRISMKNLLNHPWIMQDYNYPVEWQSKNPFILHDDDCVTLSVHHRNNROT 300
QY 301 MEDLISLWQYDHLTATYLLILLAKKARGKPVRLRLSSFCGQASATPFTDIKSNNWSLEDV 360
DB 301 MEDLISLWQYDHLTATYLLILLAKKARGKPVRLRLSSFCGQASATPFTDIKSNNWSLEDV 360
QY 361 TASDKNYYVAGLIDYDWCEDDLSTGAATPRTSQFTKYWTESNGVESKSLTPALCRTPANKL 420
DB 361 TASDKNYYVAGLIDYDWCEDDLSTGAATPRTSQFTKYWTESNGVESKSLTPALCRTPANKL 420
QY 421 KKNENYVTPKSAVKNEEYFMPEPKTPVNKNQHKREILTPNRYTTPSKARNQCLKETPI 480
DB 421 KKNENYVTPKSAVKNEEYFMPEPKTPVNKNQHKREILTPNRYTTPSKARNQCLKETPI 480
QY 481 KIPVNSTGTDKMTGTVISPERCRSVELDLNOAHMEETPKRGAKVFGSLERGLDKVITV 540
DB 481 KIPVNSTGTDKMTGTVISPERCRSVELDLNOAHMEETPKRGAKVFGSLERGLDKVITV 540
QY 541 LTRSKRKGSGARDGPRRLKLYHNVTTTRELVNPDLQLLNEIMSILPKKHVDVFOQGYTLKCQT 600
DB 541 LTRSKRKGSGARDGPRRLKLYHNVTTTRELVNPDLQLLNEIMSILPKKHVDVFOQGYTLKCQT 600
QY 601 QSDFGKVTMQFLEVCQLOKPDVVGIRRORLKGDAWYKRLVEDIILSSCKV 651
DB 601 QSDFGKVTMQFLEVCQLOKPDVVGIRRORLKGDAWYKRLVEDIILSSCKV 651
```

RESULT 10  
US-11-177-138-10  
; Sequence 10, Application US/11177138

```
; Publication No. US20050282766A1
; GENERAL INFORMATION:
; APPLICANT: Wu, Bin
; APPLICANT: Seeley, Todd
; APPLICANT: Williams, Lewis T.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING NEOPLASTIC DISEASE USING
; FILE REFERENCE: 200130.514/PP-01623.002
; CURRENT APPLICATION NUMBER: US/11/177,138
; CURRENT FILING DATE: 2005-07-08
; PRIOR APPLICATION NUMBER: US/09/870,937
; PRIOR FILING DATE: 2001-05-30
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 10
; LENGTH: 651
; TYPE: PRT
; ORGANISM: Homo sapien
US-11-177-138-10

Query Match      100.0%; Score 3447; DB 6; Length 651;
Best Local Similarity 100.0%; Pred. No. 7.8e-251;
Matches 651; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKDYDELLKYYELHETIGTGPAKVKLACHILTGEMVAIKIMDKNTLGSDDLPRIKTEIEA 60
DB 1 MKDYDELLKYYELHETIGTGPAKVKLACHILTGEMVAIKIMDKNTLGSDDLPRIKTEIEA 60
QY 61 LKNLRHQHICQLYHVLETANKIFMWLEYCPGGELFDYIISQDRLSEETRVVFRQIVSAV 120
DB 61 LKNLRHQHICQLYHVLETANKIFMWLEYCPGGELFDYIISQDRLSEETRVVFRQIVSAV 120
QY 121 AYVHSQGYAHRDLKPNLLFDEYHKLIDFGLCAKPGKNDYHLQTCGSLAYAAPELI 180
DB 121 AYVHSQGYAHRDLKPNLLFDEYHKLIDFGLCAKPGKNDYHLQTCGSLAYAAPELI 180
QY 181 QGKSYLGSEADVWSMGILLVLMCGFLPDDDDNMALYKKIMRGKYDVPKWLSPSSILL 240
DB 181 QGKSYLGSEADVWSMGILLVLMCGFLPDDDDNMALYKKIMRGKYDVPKWLSPSSILL 240
QY 241 QQMLQVDPKKRISMKNLLNHPWIMQDYNYPVEWQSKNPFILHDDDCVTLSVHHRNNROT 300
DB 241 QQMLQVDPKKRISMKNLLNHPWIMQDYNYPVEWQSKNPFILHDDDCVTLSVHHRNNROT 300
QY 301 MEDLISLWQYDHLTATYLLILLAKKARGKPVRLRLSSFCGQASATPFTDIKSNNWSLEDV 360
DB 301 MEDLISLWQYDHLTATYLLILLAKKARGKPVRLRLSSFCGQASATPFTDIKSNNWSLEDV 360
QY 361 TASDKNYYVAGLIDYDWCEDDLSTGAATPRTSQFTKYWTESNGVESKSLTPALCRTPANKL 420
DB 361 TASDKNYYVAGLIDYDWCEDDLSTGAATPRTSQFTKYWTESNGVESKSLTPALCRTPANKL 420
QY 421 KKNENYVTPKSAVKNEEYFMPEPKTPVNKNQHKREILTPNRYTTPSKARNQCLKETPI 480
DB 421 KKNENYVTPKSAVKNEEYFMPEPKTPVNKNQHKREILTPNRYTTPSKARNQCLKETPI 480
QY 481 KIPVNSTGTDKMTGTVISPERCRSVELDLNOAHMEETPKRGAKVFGSLERGLDKVITV 540
DB 481 KIPVNSTGTDKMTGTVISPERCRSVELDLNOAHMEETPKRGAKVFGSLERGLDKVITV 540
QY 541 LTRSKRKGSGARDGPRRLKLYHNVTTTRELVNPDLQLLNEIMSILPKKHVDVFOQGYTLKCQT 600
DB 541 LTRSKRKGSGARDGPRRLKLYHNVTTTRELVNPDLQLLNEIMSILPKKHVDVFOQGYTLKCQT 600
QY 601 QSDFGKVTMQFLEVCQLOKPDVVGIRRORLKGDAWYKRLVEDIILSSCKV 651
DB 601 QSDFGKVTMQFLEVCQLOKPDVVGIRRORLKGDAWYKRLVEDIILSSCKV 651
```

RESULT 11  
US-10-425-114-54260  
; Sequence 54260, Application US/10425114  
; Publication No. US20040034888A1



; GENERAL INFORMATION:  
; APPLICANT: Liu, Jingdong  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Screen, Steven E  
; APPLICANT: Tabaska, Jack E  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53313)B  
; CURRENT APPLICATION NUMBER: US/10/425,114  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 73128  
; SEQ ID NO 54260  
; LENGTH: 656  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: Clone ID: LIB4649-035-D9\_FLI.pep  
US-10-425-114-54260

Query Match 100.0%; Score 3447; DB 4; Length 656;  
Best Local Similarity 100.0%; Pred. No. 7.9e-251;  
Matches 651; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MKDYDELLKYELHETIGTGPAKVKLACHILTGEMVAIKIMDKNTLGSGLPDKTEIEA	60
DB	6	MKDYDELLKYELHETIGTGPAKVKLACHILTGEMVAIKIMDKNTLGSGLPDKTEIEA	65
QY	61	LKNLRHQHICOLYHVLETANKIFWLEYCPGGLFDYIISODRLSEETRVVFRQIVSAV	120
DB	66	LKNLRHQHICOLYHVLETANKIFWLEYCPGGLFDYIISODRLSEETRVVFRQIVSAV	125
QY	121	AYVHSQGYAHRDLKPNELLFDEYHKLKIDFGLCAKPGKNDYHLOTCCGSLAYAAPELI	180
DB	126	AYVHSQGYAHRDLKPNELLFDEYHKLKIDFGLCAKPGKNDYHLOTCCGSLAYAAPELI	185
QY	181	QKSYLGSSEADVWSMGILLVLMCGFLPFDDNNMALYKIMRGKYDVPKWLSPSSILL	240
DB	186	QKSYLGSSEADVWSMGILLVLMCGFLPFDDNNMALYKIMRGKYDVPKWLSPSSILL	245
QY	241	QOMLQVDPKKRISMKNLNHPWIMQDYNYPVEWQSKNPFILHDDCCVTLSVHHRNROT	300
DB	246	QOMLQVDPKKRISMKNLNHPWIMQDYNYPVEWQSKNPFILHDDCCVTLSVHHRNROT	305
QY	301	MEDLISLWQYDHLTATYLLILLAKKARGKPVRLSLSSFCQASATPFTDIKSNWNSLEDV	360
DB	306	MEDLISLWQYDHLTATYLLILLAKKARGKPVRLSLSSFCQASATPFTDIKSNWNSLEDV	365
QY	361	TASDKNYVAGLIDYDWCEDDLSTGAATPRTSQFTKYWTESNGVESKSLTPALCPTPANKL	420
DB	366	TASDKNYVAGLIDYDWCEDDLSTGAATPRTSQFTKYWTESNGVESKSLTPALCPTPANKL	425
QY	421	KKENYVTPKSAVKNEEYFMPPEKTPVNKNQHKREILTPNRYTTPSKARNOCLKETPI	480
DB	426	KKENYVTPKSAVKNEEYFMPPEKTPVNKNQHKREILTPNRYTTPSKARNOCLKETPI	485
QY	481	KIPVNSTGTDKLMGTGVI SPERRCRSVELDLNOAHMEETPKRGAKVFGSLERGLDKVITV	540
DB	486	KIPVNSTGTDKLMGTGVI SPERRCRSVELDLNOAHMEETPKRGAKVFGSLERGLDKVITV	545
QY	541	LTRSKRKSARDGPRRLKHLHYNTTTRLVNPDQLNEIMSILPKGHVDFVQGYTLKCOQT	600
DB	546	LTRSKRKSARDGPRRLKHLHYNTTTRLVNPDQLNEIMSILPKGHVDFVQGYTLKCOQT	605
QY	601	QSDFGKVTWQFELEVCOLOKPDVVGIRRORLKGDAWYKRLVEDILSSCKV	651
DB	606	QSDFGKVTWQFELEVCOLOKPDVVGIRRORLKGDAWYKRLVEDILSSCKV	656

RESULT 12  
US-10-425-114-54261  
; Sequence 54261, Application US/10425114

; Publication No. US20040034888A1  
; GENERAL INFORMATION:  
; APPLICANT: Liu, Jingdong  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Screen, Steven E  
; APPLICANT: Tabaska, Jack E  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53313)B  
; CURRENT APPLICATION NUMBER: US/10/425,114  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 73128  
; SEQ ID NO 54261  
; LENGTH: 656  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: Clone ID: LIB4676-059-G1\_FLI.pep  
US-10-425-114-54261

Query Match 100.0%; Score 3447; DB 4; Length 656;  
Best Local Similarity 100.0%; Pred. No. 7.9e-251;  
Matches 651; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MKDYDELLKYELHETIGTGPAKVKLACHILTGEMVAIKIMDKNTLGSGLPDKTEIEA	60
DB	6	MKDYDELLKYELHETIGTGPAKVKLACHILTGEMVAIKIMDKNTLGSGLPDKTEIEA	65
QY	61	LKNLRHQHICOLYHVLETANKIFWLEYCPGGLFDYIISODRLSEETRVVFRQIVSAV	120
DB	66	LKNLRHQHICOLYHVLETANKIFWLEYCPGGLFDYIISODRLSEETRVVFRQIVSAV	125
QY	121	AYVHSQGYAHRDLKPNELLFDEYHKLKIDFGLCAKPGKNDYHLOTCCGSLAYAAPELI	180
DB	126	AYVHSQGYAHRDLKPNELLFDEYHKLKIDFGLCAKPGKNDYHLOTCCGSLAYAAPELI	185
QY	181	QKSYLGSSEADVWSMGILLVLMCGFLPFDDNNMALYKIMRGKYDVPKWLSPSSILL	240
DB	186	QKSYLGSSEADVWSMGILLVLMCGFLPFDDNNMALYKIMRGKYDVPKWLSPSSILL	245
QY	241	QOMLQVDPKKRISMKNLNHPWIMQDYNYPVEWQSKNPFILHDDCCVTLSVHHRNROT	300
DB	246	QOMLQVDPKKRISMKNLNHPWIMQDYNYPVEWQSKNPFILHDDCCVTLSVHHRNROT	305
QY	301	MEDLISLWQYDHLTATYLLILLAKKARGKPVRLSLSSFCQASATPFTDIKSNWNSLEDV	360
DB	306	MEDLISLWQYDHLTATYLLILLAKKARGKPVRLSLSSFCQASATPFTDIKSNWNSLEDV	365
QY	361	TASDKNYVAGLIDYDWCEDDLSTGAATPRTSQFTKYWTESNGVESKSLTPALCPTPANKL	420
DB	366	TASDKNYVAGLIDYDWCEDDLSTGAATPRTSQFTKYWTESNGVESKSLTPALCPTPANKL	425
QY	421	KKENYVTPKSAVKNEEYFMPPEKTPVNKNQHKREILTPNRYTTPSKARNOCLKETPI	480
DB	426	KKENYVTPKSAVKNEEYFMPPEKTPVNKNQHKREILTPNRYTTPSKARNOCLKETPI	485
QY	481	KIPVNSTGTDKLMGTGVI SPERRCRSVELDLNOAHMEETPKRGAKVFGSLERGLDKVITV	540
DB	486	KIPVNSTGTDKLMGTGVI SPERRCRSVELDLNOAHMEETPKRGAKVFGSLERGLDKVITV	545
QY	541	LTRSKRKSARDGPRRLKHLHYNTTTRLVNPDQLNEIMSILPKGHVDFVQGYTLKCOQT	600
DB	546	LTRSKRKSARDGPRRLKHLHYNTTTRLVNPDQLNEIMSILPKGHVDFVQGYTLKCOQT	605
QY	601	QSDFGKVTWQFELEVCOLOKPDVVGIRRORLKGDAWYKRLVEDILSSCKV	651
DB	606	QSDFGKVTWQFELEVCOLOKPDVVGIRRORLKGDAWYKRLVEDILSSCKV	656

RESULT 13  
US-10-425-114-54262

```
; Sequence 54262, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(5313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 54262
; LENGTH: 656
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB4119-174-G9_FLI.pep
US-10-425-114-54262

Query Match      100.0%; Score 3447; DB 4; Length 656;
Best Local Similarity 100.0%; Pred. No. 7.9e-251;
Matches 651; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 MKDYDELLKYYELHETIGTGGAFAKVKLACHILTGEMVAIKIMDKNTLGSDDLPRIKTEISA 60
DB      6 MKDYDELLKYYELHETIGTGGAFAKVKLACHILTGEMVAIKIMDKNTLGSDDLPRIKTEISA 65

QY      61 LKNLRHQHICQLYHVLETTANKIFMWLEVCYCGGSLFDYIISQRLSBEETRVVFRQIVSAV 120
DB      66 LKNLRHQHICQLYHVLETTANKIFMWLEVCYCGGSLFDYIISQRLSBEETRVVFRQIVSAV 125

QY      121 AYVHSQGYAHRDLKPENLIFDEYHKLKIDFGLCAKPGKNDYHLQTCGSLAYAAPELI 180
DB      126 AYVHSQGYAHRDLKPENLIFDEYHKLKIDFGLCAKPGKNDYHLQTCGSLAYAAPELI 185

QY      181 QGKSYLGSEADVWSMGILLVLMCGFLPDDDDNMALYKIMRGKYDVPKWLSP 240
DB      186 QGKSYLGSEADVWSMGILLVLMCGFLPDDDDNMALYKIMRGKYDVPKWLSP 245

QY      241 QQMLQVDPKKRISMKNLNHPWIMQDYNYPVEWQSKNPFTHLDDDCVTBLSVHHRNNROT 300
DB      246 QQMLQVDPKKRISMKNLNHPWIMQDYNYPVEWQSKNPFTHLDDDCVTBLSVHHRNNROT 305

QY      301 MEDLISLQYDHLTATYLLLLAKKARGKPVRLRLSLSFSCGQASATPFTDIKNNNWSLEDV 360
DB      306 MEDLISLQYDHLTATYLLLLAKKARGKPVRLRLSLSFSCGQASATPFTDIKNNNWSLEDV 365

QY      361 TASDKNYVAGLIDYDHCEDDLSTGAATPRTSQFTKYWTESNGVESKSLTPALCRTPANKL 420
DB      366 TASDKNYVAGLIDYDHCEDDLSTGAATPRTSQFTKYWTESNGVESKSLTPALCRTPANKL 425

QY      421 KKNENYVTPKSAVKNBEYFMFPBKTPVKNQKHREILTPNRYTTPSKARNQCLKETPI 480
DB      426 KKNENYVTPKSAVKNBEYFMFPBKTPVKNQKHREILTPNRYTTPSKARNQCLKETPI 485

QY      481 KIPVNSTGTDKLMGTGISPERRCSRVELDLNQAHMEETPKRGAKVFGSLERGLDKVITV 540
DB      486 KIPVNSTGTDKLMGTGISPERRCSRVELDLNQAHMEETPKRGAKVFGSLERGLDKVITV 545

QY      541 LTRSKRGSGARDGPRILKLHYNVTTTLNPDQLNEIMSLPKKHVDVFOKGTYLKCQT 600
DB      546 LTRSKRGSGARDGPRILKLHYNVTTTLNPDQLNEIMSLPKKHVDVFOKGTYLKCQT 605

QY      601 QSDFGKVTMQFELVCOQKPDVVGIRRRQRLKGDWVYKRLVEDILSSCKV 651
DB      606 QSDFGKVTMQFELVCOQKPDVVGIRRRQRLKGDWVYKRLVEDILSSCKV 656
```

. RESULT 14

```
US-09-764-875-906
; Sequence 906, Application US/09764875
; Publication No. US20040018969A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PJ202
; CURRENT APPLICATION NUMBER: US/09/764,875
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1249
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 906
; LENGTH: 251
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (206)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (219)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (234)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (250)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-764-875-906

Query Match      35.4%; Score 1220; DB 3; Length 251;
Best Local Similarity 98.3%; Pred. No. 1.5e-83;
Matches 230; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY      1 MKDYDELLKYYELHETIGTGGAFAKVKLACHILTGEMVAIKIMDKNTLGSDDLPRIKTEISA 60
DB      6 MKDYDELLKYYELHETIGTGGAFAKVKLACHILTGEMVAIKIMDKNTLGSDDLPRIKTEISA 65

QY      61 LKNLRHQHICQLYHVLETTANKIFMWLEVCYCGGSLFDYIISQRLSBEETRVVFRQIVSAV 120
DB      66 LKNLRHQHICQLYHVLETTANKIFMWLEVCYCGGSLFDYIISQRLSBEETRVVFRQIVSAV 125

QY      121 AYVHSQGYAHRDLKPENLIFDEYHKLKIDFGLCAKPGKNDYHLQTCGSLAYAAPELI 180
DB      126 AYVHSQGYAHRDLKPENLIFDEYHKLKIDFGLCAKPGKNDYHLQTCGSLAYAAPELI 185

QY      181 QGKSYLGSEADVWSMGILLVLMCGFLPDDDDNMALYKIMRGKYDVPKWLSP 234
DB      186 QGKSYLGSEADVWSMGILLVLMCGFLPDDDDNMALYKIMRGKYDVGKXWLSP 239

RESULT 15
US-10-106-698-5313
; Sequence 5313, Application US/10106698
; Publication No. US20030109690A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Colon and Colon Cancer Associated Polynucleotides and Polypeptides
; FILE REFERENCE: PA005P1
; CURRENT APPLICATION NUMBER: US/10/106,698
; CURRENT FILING DATE: 2002-03-27
; PRIOR APPLICATION NUMBER: PCT/US00/26524
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US 60/157,137
; PRIOR FILING DATE: 1999-09-29
; PRIOR APPLICATION NUMBER: US 60/163,280
; PRIOR FILING DATE: 1999-11-03
; NUMBER OF SEQ ID NOS: 8564
; SOFTWARE: PatentIn Ver. 3.0
; SEQ ID NO 5313
; LENGTH: 272
; TYPE: PRT
```

```
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (220)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-10-106-698-5313

Query Match      34.1%; Score 1176; DB 4; Length 272;
Best Local Similarity 84.1%; Pred. No. 3.4e-80;
Matches 227; Conservative 8; Mismatches 35; Indels 0; Gaps 0;

QY 285 DCVTELSVHRRNRQTWEDLISLWQYDHLTATYLLLLAKKARGKPVRLRLSSPSCGOASA 344
Db 1 DCVTELSVHRRNRQTWEDLISLWQYDHLTATYLLLLAKKARGKPVRLRLSSPSCGOASA 60

QY 345 TPFTDIKSNNWSLEDVTASDKNYVAGLIDYDWCEDDLSTGAAATPRTSQFTKYWTESNGVE 404
Db 61 TPFTDIKSNNWSLEDVTASDKNYVAGLIDYDWCEDDLSTGAAATPRTSQFTKYWTESNGVE 120

QY 405 SKSLTPALCRTPANKLKNKENVTTPKSAVNKEEYFMFPEPKTPVNKNQHKREILTTPNRY 464
Db 121 SKSLTPALCRTPANKLKNKENVTTPKSAVNKEEYFMFPEPKTPVNKNQHKREILTTPNRY 180

QY 465 TTESKARNOCLETPIKIPVNSTGDKLMTGVLSPERRCRSVELDLNQAHMEETPKRKGA 524
Db 181 TTESKARNOCLETPIKIPVNSTGDKLMTGVLSPERRCRSVELDLNQAHMEETPKRKGA 240

QY 525 KVEGSLERGLDKVITVLTSTRSKRKSARDGP 554
Db 241 KCLGALKGGWIRLSLCSGPGAKGRVLPETGP 270
```

Search completed: November 21, 2006, 14:18:29  
Job time : 190 secs

**This Page Blank (uspto)**

GenCore version 5.1.9  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: November 21, 2006, 14:15:30 ; Search time 44 Seconds  
(without alignments)  
1264.753 Million cell updates/sec

Title: US-10-656-598-2  
Perfect score: 3447  
Sequence: 1 MKDYDELLKYYELHETIGT.....KGDWYVKRLVEDILSSCKV 651

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 330775 seqs, 85482512 residues

Total number of hits satisfying chosen parameters: 330775

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications AA New.\*

- 1: /EMC Celerra\_SIDS3/ptodata/2/pubpaa/US09\_NEW\_PUB.pap.\*
- 2: /EMC Celerra\_SIDS3/ptodata/2/pubpaa/US06\_NEW\_PUB.pap.\*
- 3: /EMC Celerra\_SIDS3/ptodata/2/pubpaa/US07\_NEW\_PUB.pap.\*
- 4: /EMC Celerra\_SIDS3/ptodata/2/pubpaa/US08\_NEW\_PUB.pap.\*
- 5: /EMC Celerra\_SIDS3/ptodata/2/pubpaa/PCT\_NEW\_PUB.pap.\*
- 6: /EMC Celerra\_SIDS3/ptodata/2/pubpaa/US10\_NEW\_PUB.pap.\*
- 7: /EMC Celerra\_SIDS3/ptodata/2/pubpaa/US11\_NEW\_PUB.pap.\*
- 8: /EMC Celerra\_SIDS3/ptodata/2/pubpaa/US60\_NEW\_PUB.pap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	788.5	22.9	729	US-11-377-316-154	Sequence 154, App
2	672	19.5	550	US-11-145-471-1	Sequence 1, Appl
3	672	19.5	550	US-11-145-471-23	Sequence 23, Appl
4	669	19.4	548	US-11-145-471-12	Sequence 12, Appl
5	669	19.4	548	US-11-145-471-18	Sequence 18, Appl
6	648.5	18.8	505	US-10-449-902-45099	Sequence 45099, A
7	647	18.8	461	US-10-449-902-43486	Sequence 43486, A
8	647	18.8	552	US-11-145-471-24	Sequence 24, Appl
9	644	18.7	552	US-11-145-471-36	Sequence 36, Appl
10	644	18.7	552	US-11-145-471-37	Sequence 37, Appl
11	640.5	18.6	509	US-10-449-902-50922	Sequence 50922, A
12	640.5	18.6	509	US-10-449-902-53492	Sequence 53492, A
13	637.5	18.5	545	US-10-449-902-47226	Sequence 47226, A
14	637	18.5	552	US-10-449-902-53399	Sequence 53399, A
15	636.5	18.5	826	US-10-449-902-41270	Sequence 41270, A
16	636.5	18.5	826	US-10-449-902-41270	Sequence 41270, A
17	635	18.4	372	US-10-953-349-10927	Sequence 10927, A
18	633	18.4	454	US-10-449-902-29700	Sequence 29700, A
19	631.5	18.3	494	US-11-056-3558-80926	Sequence 80926, A
20	631	18.3	626	US-11-145-471-45	Sequence 45, Appl
21	627	18.2	562	US-11-145-471-35	Sequence 35, Appl
22	627	18.2	624	US-11-145-471-46	Sequence 46, Appl
23	626.5	18.2	422	US-10-953-349-17556	Sequence 17556, A
24	626.5	18.2	794	US-11-320-072-74	Sequence 74, Appl
25	626	18.2	503	US-11-174-307B-4802	Sequence 4802, Ap

ALIGNMENTS

RESULT 1

US-11-377-316-154  
; Sequence 154, Application US/11377316  
; Publication No. US20060234344A1  
; GENERAL INFORMATION:  
; APPLICANT: FLOWMAN, GREGORY D.  
; APPLICANT: MARTINEZ, RICARDO  
; APPLICANT: WHYTE, DAVID  
; APPLICANT: SUDERSANAM, SUCHA  
; TITLE OF INVENTION: PROTEIN KINASES  
; FILE REFERENCE: 038602/1273  
; CURRENT APPLICATION NUMBER: US/11/377,316  
; CURRENT FILING DATE: 2006-03-17  
; NUMBER OF SEQ ID NOS: 269  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 154  
; LENGTH: 729  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-11-377-316-154

Query Match	22.9%	Score 788.5;	DB 7;	Length 729;
Best Local Similarity	30.0%	Pred. No. 3.7e-34;		
Matches	216;	Conservative 124;	Mismatches 256;	Indels 123; Gaps 20;
QY	11	YELHETIGTGSPAKVKLACHILTGEMVAIKIMDKNTLG-SDLPRIKTEIALKNLRHQHI	69	
Db	56	YRLAKTIGKGNFAKVKLARHILTGREVAIKIIDKTQNPSTLQKLPREVRIMKLNHPNI	115	
QY	70	COLYHVEITANKIPWLEVCPCGGLFDYIIISODRLSEBETRVVFRQIVSAVAVHSQGYA	129	
Db	116	VKLFEVETETKLYLTIMYASGEVFDYLVAGRMKEKRSKFRQIVSAVQVCHOKRIV	175	
QY	130	HRDLKPNLLFDYHKKLIDFGLCAK-PKGNKDVHQLTCCGSLAFAAPELIQKSYLGS	188	
Db	176	HRDLKAENLLDADNMNIADFGSNEFTVGK---LDTFCGSPPYAAAPELFGKKYDGP	232	
QY	189	EADVMSGILLVLMCGFLPFDDNNVMAKYKIMRGKYDVPKWLSPSSILLLOOMLQVDP	248	
Db	233	EVDVMSGLVILTVLSGSLPFDGQNLKELRERVLGRKYRIPFYMSDTCENLKRFLVLNP	292	
QY	249	KGRISKNNLHPWIMQDYNYPVWQSKNPFTH----LDDCVTELSEVHRNNRQTMEDL	304	
Db	293	IKRGTLQIMQDRIW----NAGHEEDLPFFPELPDIDSKRIDIMVGMGYSQEBIQS	348	
QY	305	ISLWQYDHLTATYLLLLAKKARGKPVRLRLSFSQCSQASATPPTDIKSNW-----	355	
Db	349	LSRMKYDEITATY-LLGLRKSSELDASDSSSSNLSLAKVRPSSDLNNSTGSPPHKVOR	407	

QY 356 -----SLEDVTASDKNYVAGLIDYDWCEDDLST----- 383  
Db 408 SVSSSQKORYSDHAGPAISVAYPKBSQTSADGDLKEDGSSRKSSGSAVGGKIAP 467  
QY 384 -----GAATP-----RTSOFTKYWTESNGVSKSLTPALCRTPANKLKNKENVYT 428  
Db 468 ASPMLGNASPNKADIPERKSSKTSVPSSNTASGWTTRNTYVCSERTTADR----- 518  
QY 429 PKSAVN-EEYFPPPKTPVKNQHKREILITTPNRYTTPSKARNQCLKETPIKIPVNST 487  
Db 519 -HSVIQNGKNSPTIPQRTPV-ASTHSISSAATPDR-----IRFP---R 557  
QY 488 GTDKLMTGVISP-ERRC-----RSVELDLNOAHMEETPKRKGAKVFGSLERGLDKVIT 539  
Db 558 GTASRSTFHQOPERRTATYNGPPAPSPLSHEATPLSQTSRSGSTNLF-----SKLNS 610  
QY 540 VLTRSK-----RKSGARDG-PRBLKLHYNVTTTLVNPQDLNNEIILPKKHVDVQK- 592  
Db 611 KLTRSRNVSAEQDENKEAKPRSLRFTWSMKTTSMDPGDMREIRKVLNANNCDYEORE 670  
QY 593 GYTLKCOQSDFGKVTMOPLEVCOLQKPDVVGIRQRLKGDWVVKLVEDILSSCKV 651  
Db 671 RFLLCVHGDGHAENLVQWEMEYCKPLRLSLNGVRFKRISGTSGIAPKNTASKIANELKL 729  
RESULT 2  
US-11-145-471-1  
; Sequence 1, Application US/11145471  
; Publication No. US20060147947A1  
; GENERAL INFORMATION:  
; APPLICANT: Apfeld, Javier  
; APPLICANT: O'Conner, Gregory  
; TITLE OF INVENTION: AMPK Pathway Components  
; FILE REFERENCE: 13407-0570011  
; CURRENT APPLICATION NUMBER: US/11/145,471  
; CURRENT FILING DATE: 2005-06-03  
; PRIOR APPLICATION NUMBER: PCT/US/2003/38628  
; PRIOR FILING DATE: 2003-12-04  
; PRIOR APPLICATION NUMBER: US 60/430,804  
; PRIOR FILING DATE: 2002-12-04  
; PRIOR APPLICATION NUMBER: US 60/488,261  
; PRIOR FILING DATE: 2003-07-18  
; PRIOR APPLICATION NUMBER: 60/578,804  
; PRIOR FILING DATE: 2004-06-10  
; NUMBER OF SEQ ID NOS: 56  
; SOFTWARE: Patent in version 3.2  
; SEQ ID NO 1  
; LENGTH: 550  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-11-145-471-1  
Query Match 19.5%; Score 672; DB 7; Length 550;  
Best Local Similarity 33.1%; Pred. No. 3.6e-28;  
Matches 177; Conservative 80; Mismatches 161; Indels 116; Gaps 16;  
QY 10 YYELHETIGGGPAKVKLACHILTGEMVAIKMDKNTLGS--DLPRIKTEIEALKNLRHQ 67  
Db 17 HYTLGDTLGVGTGKVKVGHKLTGHKVAVKILNRQKIRSLDVVGKIRREIQNLKLFRRHP 76  
QY 68 HICQLYHVLGTANKI FFWLEPCGGELFDYIIISQDRLSEETRVVFRQIVSAVAYVHSQ 127  
Db 77 HIILYQVISTPDSIFMWMYVSGGELFDYICNGRLDEKESRRLFQIILSGVDYCHRM 136  
QY 128 YAHRLDPENLLFDEYHKLIDFGLCAKPKGNKYHLOTCCGSLAYAAPELIQGSYLG 187  
Db 137 VVHRDLKPNVLLDAHNAKIADPGLSNMDSGE--FLRTSCGSPNYAAPEVISGRLYAG 194  
QY 188 SEADVWSMGLIYVLMCGFLPDDDDNVMALYKKIMRGKYDVPKWLSPSLLILQOMLQVD 247  
Db 195 PEVDIWSGVLVALLCGTLFPDDDDHVPFTLFFKICDGIFFYTPQYLPNSVLSLKHMLQVD 254  
QY 248 PKKRISMKNLNHPWIMQD---YNYPVWQSKNPFILHDDDCVTEL-----S 291  
Db 255 PMKRAISKIDIREHEWFKQDLPKLPDPDPSSYSTMI--DDEALKEVCEKFECESEEVLS 312  
QY 292 VHRN-----NRQTMEDL-----ISLWQYDHLT-----ATYLLLL 321

QY 248 PKKRISMKNLNHPWIMQD---YNYPVWQSKNPFILHDDDCVTEL-----S 291  
Db 255 PMKRAISKIDIREHEWFKQDLPKLPDPDPSSYSTMI--DDEALKEVCEKFECESEEVLS 312  
QY 292 VHRN-----NRQTMEDL-----ISLWQYDHLT-----ATYLLLL 321  
Db 313 LYNRNHHQDPLAVAYHLIIDNRIMNEAKDFYIATSPDSPDSDLDDHHLTRHPRPERVPLVAE 372  
QY 322 AKKAR-----GKPVRLRLSPSCGQASATPTDIKSN-----NW----- 355  
Db 373 TPRARHTLDELNPQSKHQGVKAKWHLGIQSQRENDIMAEVCAIKQDLDYEMKVNPY 432  
QY 356 -----SLEDVTASDKNYVAGLIDYDWCEDDLST-----TGAATP----- 388  
Db 433 YLRVRKKNPVTSTYKMSLQLYQVDSRTY---LLDFRSIDDEITEAKSGTATPQRSQSVS 489  
QY 389 --RTSOFTKYWTESNGVESK-SLTPALCRTPANKLKNKENVYTPKSAVKNBEYF 439  
Db 490 NYRSQORSDDAEAQGSSEVSLTSSV-----TSLDSSPVDLTPRPGSHTEIFF 538  
RESULT 3  
US-11-145-471-23  
; Sequence 23, Application US/11145471  
; Publication No. US20060147947A1  
; GENERAL INFORMATION:  
; APPLICANT: Apfeld, Javier  
; APPLICANT: O'Conner, Gregory  
; TITLE OF INVENTION: AMPK Pathway Components  
; FILE REFERENCE: 13407-0570011  
; CURRENT APPLICATION NUMBER: US/11/145,471  
; CURRENT FILING DATE: 2005-06-03  
; PRIOR APPLICATION NUMBER: PCT/US/2003/38628  
; PRIOR FILING DATE: 2003-12-04  
; PRIOR APPLICATION NUMBER: US 60/430,804  
; PRIOR FILING DATE: 2002-12-04  
; PRIOR APPLICATION NUMBER: US 60/488,261  
; PRIOR FILING DATE: 2003-07-18  
; PRIOR APPLICATION NUMBER: 60/578,804  
; PRIOR FILING DATE: 2004-06-10  
; NUMBER OF SEQ ID NOS: 56  
; SOFTWARE: Patent in version 3.2  
; SEQ ID NO 23  
; LENGTH: 550  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-11-145-471-23  
Query Match 19.5%; Score 672; DB 7; Length 550;  
Best Local Similarity 33.1%; Pred. No. 3.6e-28;  
Matches 177; Conservative 80; Mismatches 161; Indels 116; Gaps 16;  
QY 10 YYELHETIGGGPAKVKLACHILTGEMVAIKMDKNTLGS--DLPRIKTEIEALKNLRHQ 67  
Db 17 HYTLGDTLGVGTGKVKVGHKLTGHKVAVKILNRQKIRSLDVVGKIRREIQNLKLFRRHP 76  
QY 68 HICQLYHVLGTANKI FFWLEPCGGELFDYIIISQDRLSEETRVVFRQIVSAVAYVHSQ 127  
Db 77 HIILYQVISTPDSIFMWMYVSGGELFDYICNGRLDEKESRRLFQIILSGVDYCHRM 136  
QY 128 YAHRLDPENLLFDEYHKLIDFGLCAKPKGNKYHLOTCCGSLAYAAPELIQGSYLG 187  
Db 137 VVHRDLKPNVLLDAHNAKIADPGLSNMDSGE--FLRTSCGSPNYAAPEVISGRLYAG 194  
QY 188 SEADVWSMGLIYVLMCGFLPDDDDNVMALYKKIMRGKYDVPKWLSPSLLILQOMLQVD 247  
Db 195 PEVDIWSGVLVALLCGTLFPDDDDHVPFTLFFKICDGIFFYTPQYLPNSVLSLKHMLQVD 254  
QY 248 PKKRISMKNLNHPWIMQD---YNYPVWQSKNPFILHDDDCVTEL-----S 291  
Db 255 PMKRAISKIDIREHEWFKQDLPKLPDPDPSSYSTMI--DDEALKEVCEKFECESEEVLS 312  
QY 292 VHRN-----NRQTMEDL-----ISLWQYDHLT-----ATYLLLL 321

Db 313 LYNRNHQDPLAVAYHLIIDNRIMNEAKOFYLATSPDSDLDHHLTRPHERVPELVAE 372  
QY 322 AKKAR-----GKPVRLRLSSFCGQASATPPTDIKSN-----NW-----355  
Db 373 TPRARHTLDELNPQKSHQGVRAKWHGLGRSQRPNIDMAEVCRAIKQLDYEWKVNPFY 432  
QY 356 -----SLEDVTASDKNYVAGLIDYDWCEDDLS-----TGAATP-----388  
Db 433 YLRVRKRNPNVTSTYKMSLQLYQVDSRTY---LLDPRSIDDEITEAKSGTATPQRSGSVS 489  
QY 389 --RTSQFTKYWTESNGVESK-SLTPALCRTPANKLKNKENVYTPKSAVKNEEYF 439  
Db 490 NYRSQRSDSDAEQKSGSEVSLTSSV-----TSLDSSPVDLTPRGSHTIERFF 538

RESULT 4

US-11-145-471-12  
; Sequence 12, Application US/11145471  
; Publication No. US20060147947A1  
; GENERAL INFORMATION:  
; APPLICANT: Apfeld, Javier  
; APPLICANT: O'Conner, Gregory  
; TITLE OF INVENTION: AMPK Pathway Components  
; FILE REFERENCE: 13407-0570011  
; CURRENT APPLICATION NUMBER: US/11/145,471  
; PRIOR FILING DATE: 2005-06-03  
; PRIOR APPLICATION NUMBER: PCT/US/2003/38628  
; PRIOR FILING DATE: 2003-12-04  
; PRIOR APPLICATION NUMBER: US 60/430,804  
; PRIOR FILING DATE: 2002-12-04  
; PRIOR APPLICATION NUMBER: 60/488,261  
; PRIOR FILING DATE: 2003-07-18  
; PRIOR APPLICATION NUMBER: US 60/578,804  
; PRIOR FILING DATE: 2004-06-10  
; NUMBER OF SEQ ID NOS: 56  
; SOFTWARE: Patent in version 3.2  
; SEQ ID NO 12  
; LENGTH: 548  
; TYPE: PRT  
; ORGANISM: Rattus norvegicus  
US-11-145-471-12

Query Match 19.4%; Score 669; DB 7; Length 548;  
Best Local Similarity 35.9%; Pred. No. 5.1e-28;  
Matches 165; Conservative 73; Mismatches 145; Indels 76; Gaps 12;  
QY 10 YYELHETIGTGFAKVKLACHILTGEMVAIKMDKNTLGS--DLPRKTEIEALKNLRHQ 67  
Db 15 HYTLGDTLGVGTGKVKVKGHELTGHKVAVKILNRQKIRSLDVGVGKIRREIQNLKLFHRP 74  
QY 68 HICQLYHVLLETANKIPMWLEPCPGGLFDYIIISQDRLSSEETRVVPRQIVSAVAYVHSQ 127  
Db 75 HIILKLVQVISTPDSIPMWMEYVSGGELFDYICKNGRLDEKESRRLFQQILSGVDYCHRM 134  
QY 128 YAHRLDKPENLLDXYHKLKLDIFGLCAKPKGNKYHLOTCCGSLAYAAPELIQKSYLG 187  
Db 135 VVHRDLKPNVLLDAHNAKIADFGLSNMMSDGE--FLRTSCGSPNYAAPEVISGRLYAG 192  
QY 188 SEADVMSMGIILLYVLMCGFLPFDDNNVMALYKKIMRGKYDVPKWLSPSSILLQOQLQVD 247  
Db 193 PEVDIWSGVIYALLCGTLFPDHDHVPPTLFKKICDGIFFYTPQYLNPSVISLLKHLQVD 252  
QY 248 PKKRISMKVLLNHPWIMQD---YNPVQSKNPFHLLDDCCVTEL-----S 291  
Db 253 PMKRATIKDIRHEWFKQDLPKYLFPEDPSYSMTI--DDEALKVCEKFECESEEVLS 310  
QY 292 VVHRN-----NRQTMEDL-----ISLWQYDHLT-----ATYLLLL 321  
Db 311 LYNRNHQDPLAVAYHLIIDNRIMNEAKOFYLATSPDSDLDHHLTRPHERVPELVAE 370  
QY 322 AKKAR-----GKPVRLRLSSFCGQASATPPTDIKSNWNSLEDVTASDKNYVAGL 371

Db 371 TPRARHTLDELNPQKSHQGVRAKWHGLGRSQRPNIDMAEVCRAIKQ-----419  
QY 372 IDYDWCEDD---LSTGAATPRTSQFTKYWTESNGVESKS 407  
Db 420 LDYEWKVNPNPYLRVRKRNPNVTSTFSKMSLQLYQVDSRT 458

RESULT 5

US-11-145-471-18  
; Sequence 18, Application US/11145471  
; Publication No. US20060147947A1  
; GENERAL INFORMATION:  
; APPLICANT: Apfeld, Javier  
; APPLICANT: O'Conner, Gregory  
; TITLE OF INVENTION: AMPK Pathway Components  
; FILE REFERENCE: 13407-0570011  
; CURRENT APPLICATION NUMBER: US/11/145,471  
; CURRENT FILING DATE: 2005-06-03  
; PRIOR APPLICATION NUMBER: PCT/US/2003/38628  
; PRIOR FILING DATE: 2003-12-04  
; PRIOR APPLICATION NUMBER: US 60/430,804  
; PRIOR FILING DATE: 2002-12-04  
; PRIOR APPLICATION NUMBER: US 60/488,261  
; PRIOR FILING DATE: 2003-07-18  
; PRIOR APPLICATION NUMBER: 60/578,804  
; PRIOR FILING DATE: 2004-06-10  
; NUMBER OF SEQ ID NOS: 56  
; SOFTWARE: Patent in version 3.2  
; SEQ ID NO 18  
; LENGTH: 548  
; TYPE: PRT  
; ORGANISM: Rattus norvegicus  
US-11-145-471-18

Query Match 19.4%; Score 669; DB 7; Length 548;  
Best Local Similarity 35.9%; Pred. No. 5.1e-28;  
Matches 165; Conservative 73; Mismatches 145; Indels 76; Gaps 12;  
QY 10 YYELHETIGTGFAKVKLACHILTGEMVAIKMDKNTLGS--DLPRKTEIEALKNLRHQ 67  
Db 15 HYTLGDTLGVGTGKVKVKGHELTGHKVAVKILNRQKIRSLDVGVGKIRREIQNLKLFHRP 74  
QY 68 HICQLYHVLLETANKIPMWLEPCPGGLFDYIIISQDRLSSEETRVVPRQIVSAVAYVHSQ 127  
Db 75 HIILKLVQVISTPDSIPMWMEYVSGGELFDYICKNGRLDEKESRRLFQQILSGVDYCHRM 134  
QY 128 YAHRLDKPENLLDXYHKLKLDIFGLCAKPKGNKYHLOTCCGSLAYAAPELIQKSYLG 187  
Db 135 VVHRDLKPNVLLDAHNAKIADFGLSNMMSDGE--FLRTSCGSPNYAAPEVISGRLYAG 192  
QY 188 SEADVMSMGIILLYVLMCGFLPFDDNNVMALYKKIMRGKYDVPKWLSPSSILLQOQLQVD 247  
Db 193 PEVDIWSGVIYALLCGTLFPDHDHVPPTLFKKICDGIFFYTPQYLNPSVISLLKHLQVD 252  
QY 248 PKKRISMKVLLNHPWIMQD---YNPVQSKNPFHLLDDCCVTEL-----S 291  
Db 253 PMKRATIKDIRHEWFKQDLPKYLFPEDPSYSMTI--DDEALKVCEKFECESEEVLS 310  
QY 292 VVHRN-----NRQTMEDL-----ISLWQYDHLT-----ATYLLLL 321  
Db 311 LYNRNHQDPLAVAYHLIIDNRIMNEAKOFYLATSPDSDLDHHLTRPHERVPELVAE 370  
QY 322 AKKAR-----GKPVRLRLSSFCGQASATPPTDIKSNWNSLEDVTASDKNYVAGL 371  
Db 371 TPRARHTLDELNPQKSHQGVRAKWHGLGRSQRPNIDMAEVCRAIKQ-----419  
QY 372 IDYDWCEDD---LSTGAATPRTSQFTKYWTESNGVESKS 407  
Db 420 LDYEWKVNPNPYLRVRKRNPNVTSTFSKMSLQLYQVDSRT 458

RESULT 6

US-10-449-902-45099

```
; Sequence 45099, Application US/10449902
; Publication No. US20060123505A1
; GENERAL INFORMATION:
; APPLICANT: National Institute of Agrobiological Sciences.
; APPLICANT: Bio-oriented Technology Research Advancement Institution.
; APPLICANT: The Institute of Physical and Chemical Research.
; APPLICANT: Foundation for Advancement of International Science.
; TITLE OF INVENTION: FULL-LENGTH PLANT CDNA AND USES THEREOF
; FILE REFERENCE: MOA-A0205Y1-US
; CURRENT APPLICATION NUMBER: US/10/449,902
; CURRENT FILING DATE: 2003-05-29
; PRIOR APPLICATION NUMBER: JP 2002-203269
; PRIOR FILING DATE: 2002-05-30
; PRIOR APPLICATION NUMBER: JP 2002-383870
; PRIOR FILING DATE: 2002-12-11
; NUMBER OF SEQ ID NOS: 56791
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 45099
; LENGTH: 505
; TYPE: PRT
; ORGANISM: Oryza sativa
; US-10-449-902-45099

Query Match      18.8%; Score 648.5; DB 6; Length 505;
Best Local Similarity 39.8%; Pred. No. 5.6e-27;
Matches 155; Conservative 57; Mismatches 138; Indels 39; Gaps 8;

QY 11 YELHETIGTGGAFAKVKLACHILTGEMVAIKIMDKNTLGS--DLPRIKTEALKNLRHOH 68
DB 14 YRIGKTGIGSGFGKVKIAEHILTGHKVAIKILNRKIKSMEMEEKVREIKILRFLMHPH 73

QY 69 ICOLYHVLETANKIFMWLEPCGGELEFDYIISQDRLSEETRVVFRQIVSAVYVHSQGY 128
DB 74 IIRLYEVIDTPADIVYVMEYVKGSELEFDYIVKEGRQEEARFFQOIISGVEYCHRMV 133

QY 129 AHRDLKPENLLFDEYHKLIIDFGLCAKPKGNKYD--LQTCGSLAYAAPELIQSGSYLG 187
DB 134 VHRDLKPENLLSKCNVKADEFL---SNVWRDGHFLKTSCTGSPNYAAPEVSGKLYAG 190

QY 188 SEADVWSGILLVLMCGELPFDDNNMALKYKIMRGKYDVPKWLSPSSILLQQMLQVLD 247
DB 191 PEVDVWSGILLVALLCGTLPFDENIPNLFKIKGIIYTLPSHLPLARDLLPRMLVVD 250

QY 248 PKKRISKNLNNHPWI---MDYNYPVWQSKNPFILHDDDCVTLSVHRRNNRQTMEDL 304
DB 251 PMKRIITREHQWFTVGLPRYLAVPPPTAQVKKLDDDETLDVINGMFGDKNQLIESL 310

QY 305 ISLWQVDHLTATYLLLLAKKAKGKPVRLRLSSFSGQ-----ASATP-----FT 348
DB 311 HKRLQNEATVAYYLL-----DNRLRTSGYLGAEPHESMESSLAQVTPAETPNSAT 362

QY 349 P-----IKSNWSLEDTVASDKNYVAGL 371
DB 363 DHRQGHMESPGFGLRHHFAADRRKWAJGL 391

RESULT 7
US-10-449-902-43486
; Sequence 43486, Application US/10449902
; Publication No. US20060123505A1
; GENERAL INFORMATION:
; APPLICANT: National Institute of Agrobiological Sciences.
; APPLICANT: Bio-oriented Technology Research Advancement Institution.
; APPLICANT: The Institute of Physical and Chemical Research.
; APPLICANT: Foundation for Advancement of International Science.
; TITLE OF INVENTION: FULL-LENGTH PLANT CDNA AND USES THEREOF
; FILE REFERENCE: MOA-A0205Y1-US
; CURRENT APPLICATION NUMBER: US/10/449,902
; CURRENT FILING DATE: 2003-05-29
; PRIOR APPLICATION NUMBER: JP 2002-203269
; PRIOR FILING DATE: 2002-05-30
; PRIOR APPLICATION NUMBER: JP 2002-383870
; PRIOR FILING DATE: 2002-12-11

; NUMBER OF SEQ ID NOS: 56791
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 43486
; LENGTH: 461
; TYPE: PRT
; ORGANISM: Oryza sativa
; US-10-449-902-43486

Query Match      18.8%; Score 647; DB 6; Length 461;
Best Local Similarity 46.5%; Pred. No. 6.1e-27;
Matches 132; Conservative 53; Mismatches 83; Indels 16; Gaps 5;

QY 7 LLKYVELHETIGTGGAFAKVKLACHILTGEMVAIKIMDKN---TLGSDLPRIKTEIALKN 63
DB 15 LLGRYEIGRTGLGEGNFGKVKYARHLATGAHFAIKILDRNKILSLRPD-QDIRREIGTLKL 73

QY 64 LRHQHICOLYHVLETANKIFMWLEPCGGELEFDYIISQDRLSEETRVVFRQIVSAVAV 123
DB 74 LKHPNVVRLHEVAASKTKIYMLVLEVYVNGGELFDKIAVKGKLSHEGRRLLFQQIDAVSYC 133

QY 124 HSQGYAHRDLKPENLLFDEYHKLIIDFGLCAKPK--GNKDYHLQTCGSLAYAAPELIQ 181
DB 134 HDKGVYHRDLKPENLVDRGNIKISDFGLSALPQHLGN-DGLLHTTCGSPNYIAEVLQ 192

QY 182 GKSYLESEADVWSMGILLVLMCGELPFDDNNMALKYKIMRGKYDVPKWLSPSSILLQ 241
DB 193 NRGYDGLSLDIWSCGVILYVLMVGLVLPFDDRLNLVLYQIKFKGDTQIPKWLSPSARDLLR 252

QY 242 QMLQVDPKKRISKNLNNHPWIMQDYNYPVWQSKNPFILHDDDD 285
DB 253 RILEPNMFKRIINAGIKHEHWFQDY-----TPVVVPYDDDD 287

RESULT 8
US-11-145-471-24
; Sequence 24, Application US/11145471
; Publication No. US20060147947A1
; GENERAL INFORMATION:
; APPLICANT: Apfeld, Javier
; APPLICANT: O'Conner, Gregory
; TITLE OF INVENTION: AMPK Pathway Components
; FILE REFERENCE: 13407-0570011
; CURRENT APPLICATION NUMBER: US/11/145,471
; CURRENT FILING DATE: 2005-06-03
; PRIOR APPLICATION NUMBER: PCT/US/2003/38628
; PRIOR FILING DATE: 2003-12-04
; PRIOR APPLICATION NUMBER: US 60/430,804
; PRIOR FILING DATE: 2002-12-04
; PRIOR APPLICATION NUMBER: US 60/488,261
; PRIOR FILING DATE: 2003-07-18
; PRIOR APPLICATION NUMBER: 60/578,804
; PRIOR FILING DATE: 2004-06-10
; NUMBER OF SEQ ID NOS: 56
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 24
; LENGTH: 552
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-11-145-471-24

Query Match      18.8%; Score 647; DB 7; Length 552;
Best Local Similarity 43.3%; Pred. No. 7.3e-27;
Matches 139; Conservative 51; Mismatches 115; Indels 16; Gaps 6;

QY 10 YVELHETIGTGGAFAKVKLACHILTGEMVAIKIMDKNTLGS--DLPRIKTEIALKNLRHQ 67
DB 15 HVLGDTLGVGTGKVKIGEHQLTGHKVAIKILNRKISRLDVGKIKREIQNLKLFRRHP 74

QY 68 HICQLYHVLETANKIFMWLEPCGGELEFDYIISQDRLSEETRVVFRQIVSAVYVHSQ 127
DB 75 HIILYQVISTPTDFFWMEYVSGGELFDYICGHGRVEMEARLLFQIILSAVDYCHRM 134

QY 128 YAHRLDKPENLLFDEYHKLIIDFGLCAKPKGNKYDHLQTCGSLAYAAPELIQSGSYLG 187
```



```
Db 135 VVHRDLKPENLVLDADHNAKIADFGLSNMMSDGE--FLRTSCGSPNYAAPEVISGRLYAG 192
QY 188 SEADVMSGILLVYLMCGFLPFDDNNMVALYKIMRGKYDVPKWLSPSSILLLOQMLQVD 247
Db 193 PEVDIWSGVIYALLCGTLFPDDEHVPTLFKKIRGGVFYIPEYLNRSVATLLMHMLQVD 252
QY 248 PKKRISMKNLNHPWIMQD---YNYPVWQSKNPFTH---LDDDCVTELSVHHR--NNROT 300
Db 253 PLKRATIKDIRHEWFKQDLPSYLPF-----EDPSYDANVIDDEAVKEVCEKPECTESEV 307
QY 301 MEDLISLWQYDHLTATYLLLL 321
Db 308 MNSLYSGDPQDQLAVAYHLII 328

RESULT 9
US-11-145-471-36
; Sequence 36, Application US/11145471
; Publication No. US20060147947A1
; GENERAL INFORMATION:
; APPLICANT: Apfeld, Javier
; TITLE OF INVENTION: AMPK Pathway Components
; FILE REFERENCE: 13407-0570011
; CURRENT APPLICATION NUMBER: US/11/145,471
; PRIOR FILING DATE: 2005-06-03
; PRIOR APPLICATION NUMBER: PCT/US/2003/38628
; PRIOR FILING DATE: 2003-12-04
; PRIOR FILING DATE: 2002-12-04
; PRIOR FILING DATE: 2003-07-18
; PRIOR FILING DATE: 2004-06-10
; NUMBER OF SEQ ID NOS: 56
; SOFTWARE: Patentin version 3.2
; SEQ ID NO 36
; LENGTH: 552
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-11-145-471-36

Query Match 18.7%; Score 644; DB 7; Length 552;
Best Local Similarity 43.3%; Pred. No. 1.1e-26;
Matches 139; Conservative 50; Mismatches 116; Indels 16; Gaps 6;

QY 10 YYELHETIGTGGAFAKVKLACHILTGEMVAIKMDKNTLGS--DLPRIKTEIEALKNLRHQ 67
Db 15 HYVLGDTLGVGTGKVKIGEHQLTGHKAVKILNRQIRSLDVGKIKREIQNLKLFHRP 74
QY 68 HICQLYHVLLETANKIFWVLEYCPGGBELFDYIISQDRLSSEETRVVFRQIVSAVAVYHSG 127
Db 75 HIIKLYQVISTPTDFWVMEYVSGGBELFDYICKHGRVEVEARLFFQQLSADVDYCHRM 134
QY 128 YAHRLDKPENLFPDEYHKLIDFGICAKPKGNKDYHLQTCGSLAYAAPELIQGSYLG 187
Db 135 VVHRDLKPENLVLDADHNAKIADFGLSNMMSDGE--FLRTSCGSPNYAAPEVISGRLYAG 192
QY 188 SEADVMSGILLVYLMCGFLPFDDNNMVALYKIMRGKYDVPKWLSPSSILLLOQMLQVD 247
Db 193 PEVDIWSGVIYALLCGTLFPDDEHVPTLFKKIRGGVFYIPEYLNRSVATLLMHMLQVD 252
QY 248 PKKRISMKNLNHPWIMQD---YNYPVWQSKNPFTH---LDDDCVTELSVHHR--NNROT 300
Db 253 PLKRATIKDIRHEWFKQDLPSYLPF-----EDPSYDANVIDDEAVKEVCEKPECTESEV 307
QY 301 MEDLISLWQYDHLTATYLLLL 321
Db 308 MNSLYSGDPQDQLAVAYHLII 328

RESULT 10
US-10-449-902-50922
; Sequence 50922, Application US/10449902
; Publication No. US20060123505A1
; GENERAL INFORMATION:
; APPLICANT: National Institute of Agrobiological Sciences.
; APPLICANT: Bio-oriented Technology Research Advancement Institution.
; APPLICANT: The Institute of Physical and Chemical Research.
; APPLICANT: Foundation for Advancement of International Science.
; TITLE OF INVENTION: FULL-LENGTH PLANT CDNA AND USES THEREOF
; FILE REFERENCE: MOA-A02051-US
; CURRENT APPLICATION NUMBER: US/10/449,902
; CURRENT FILING DATE: 2003-05-29
; PRIOR APPLICATION NUMBER: JP 2002-203269
; PRIOR FILING DATE: 2002-05-30
; PRIOR APPLICATION NUMBER: JP 2002-383870
; PRIOR FILING DATE: 2002-12-11
; NUMBER OF SEQ ID NOS: 56791
```

```
US-11-145-471-37
; Sequence 37, Application US/11145471
; Publication No. US20060147947A1
; GENERAL INFORMATION:
; APPLICANT: Apfeld, Javier
; TITLE OF INVENTION: AMPK Pathway Components
; FILE REFERENCE: 13407-0570011
; CURRENT APPLICATION NUMBER: US/11/145,471
; CURRENT FILING DATE: 2005-06-03
; PRIOR APPLICATION NUMBER: PCT/US/2003/38628
; PRIOR FILING DATE: 2003-12-04
; PRIOR FILING DATE: 2002-12-04
; PRIOR FILING DATE: 2003-07-18
; PRIOR FILING DATE: 2004-06-10
; NUMBER OF SEQ ID NOS: 56
; SOFTWARE: Patentin version 3.2
; SEQ ID NO 37
; LENGTH: 552
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-11-145-471-37

Query Match 18.7%; Score 644; DB 7; Length 552;
Best Local Similarity 43.3%; Pred. No. 1.1e-26;
Matches 139; Conservative 50; Mismatches 116; Indels 16; Gaps 6;

QY 10 YYELHETIGTGGAFAKVKLACHILTGEMVAIKMDKNTLGS--DLPRIKTEIEALKNLRHQ 67
Db 15 HYVLGDTLGVGTGKVKIGEHQLTGHKAVKILNRQIRSLDVGKIKREIQNLKLFHRP 74
QY 68 HICQLYHVLLETANKIFWVLEYCPGGBELFDYIISQDRLSSEETRVVFRQIVSAVAVYHSG 127
Db 75 HIIKLYQVISTPTDFWVMEYVSGGBELFDYICKHGRVEVEARLFFQQLSADVDYCHRM 134
QY 128 YAHRLDKPENLFPDEYHKLIDFGICAKPKGNKDYHLQTCGSLAYAAPELIQGSYLG 187
Db 135 VVHRDLKPENLVLDADHNAKIADFGLSNMMSDGE--FLRTSCGSPNYAAPEVISGRLYAG 192
QY 188 SEADVMSGILLVYLMCGFLPFDDNNMVALYKIMRGKYDVPKWLSPSSILLLOQMLQVD 247
Db 193 PEVDIWSGVIYALLCGTLFPDDEHVPTLFKKIRGGVFYIPEYLNRSVATLLMHMLQVD 252
QY 248 PKKRISMKNLNHPWIMQD---YNYPVWQSKNPFTH---LDDDCVTELSVHHR--NNROT 300
Db 253 PLKRATIKDIRHEWFKQDLPSYLPF-----EDPSYDANVIDDEAVKEVCEKPECTESEV 307
QY 301 MEDLISLWQYDHLTATYLLLL 321
Db 308 MNSLYSGDPQDQLAVAYHLII 328

RESULT 11
US-10-449-902-50922
; Sequence 50922, Application US/10449902
; Publication No. US20060123505A1
; GENERAL INFORMATION:
; APPLICANT: National Institute of Agrobiological Sciences.
; APPLICANT: Bio-oriented Technology Research Advancement Institution.
; APPLICANT: The Institute of Physical and Chemical Research.
; APPLICANT: Foundation for Advancement of International Science.
; TITLE OF INVENTION: FULL-LENGTH PLANT CDNA AND USES THEREOF
; FILE REFERENCE: MOA-A02051-US
; CURRENT APPLICATION NUMBER: US/10/449,902
; CURRENT FILING DATE: 2003-05-29
; PRIOR APPLICATION NUMBER: JP 2002-203269
; PRIOR FILING DATE: 2002-05-30
; PRIOR APPLICATION NUMBER: JP 2002-383870
; PRIOR FILING DATE: 2002-12-11
; NUMBER OF SEQ ID NOS: 56791
```

```

; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 50922
; LENGTH: 509
; TYPE: PRT
; ORGANISM: Oryza sativa
US-10-449-902-50922

Query Match      18.6%; Score 640.5; DB 6; Length 509;
Best Local Similarity 36.9%; Pred. No. 1.5e-26;
Matches 149; Conservative 64; Mismatches 118; Indels 73; Gaps 10;

QY      4 YDELLKYYELHETIGTGGFAKVLAACHILTGMVAIKMDKNTLGS--DLPRIKTEIAL 61
Db      10 HSEALKNYNLGRTLGIGSGFKVIAEHLKTHRVAIKILNRQRNMNEMEEKAKREIKIL 69

QY      62 KNLRHQHICOLYHVLETANKIFMWLEYCPGGELFDYIISQDRLSBEETRVFRQIVSAVA 121
Db      70 RLFPHIIRLYEVIYTPTDIYVMVEYCKFGELFDYIVEKGRLODEARRIFQOIISGYE 129

QY      122 YVHSQGYAHRDLKPENLLFDEYHKLIIDFGLCAKPKGNKDYHLOTCCGSLAYAAPELIQ 181
Db      130 YCHRNMMVVRDLKPENLLDLSKYNVKLADPGL--SNVMHDGHFLKTSCTGSPNYAAPEVIS 187

QY      182 GKSYLEGSEADVWMSGILLYVLMCGFLPFDDNNVMALYKKIMRGKYDVPKWLSPSSILLIQ 241
Db      188 GKLYAGPEVDVWMSGVILYALLCGTLPPDDENIPNLFKKIKGGIYTLPSHLSALARDLIP 247

QY      242 QMLQVDPKGRISMKNNLNHPIMQDYNYP-----VEWQSKNPFHLLDDCVTELSV 292
Db      248 RMLVVDPMKRITIREHQWF--QIRLPYLAVPPTAQOAK-----MIDED----- 294

QY      293 HHNNRQTMEDLISL-WQYDHL-----TATYLLLLAKKARG-----KP 329
Db      295 -----TLQDVVNLGVEKDHVCESLRNRLQNEATVAYLLLDNFRFRATSGYLGADYQES 347

QY      330 VRLRLSSFSCGQASATPFTDIKSNNSLEDVTASDKNYVAGLID 373
Db      348 LERNLRFA-----SSESASNTRHYPGSSD 374

RESULT 13
US-10-449-902-47226
; Sequence 47226, Application US/10449902
; Publication No. US20060123505A1
; GENERAL INFORMATION:
; APPLICANT: National Institute of Agrobiological Sciences.
; APPLICANT: Bio-oriented Technology Research Advancement Institution.
; APPLICANT: The Institute of Physical and Chemical Research.
; APPLICANT: Foundation for Advancement of International Science.
; TITLE OF INVENTION: FULL-LENGTH PLANT CDNA AND USES THEREOF
; FILE REFERENCE: MOA-A0205Y1-US
; CURRENT APPLICATION NUMBER: US/10/449,902
; PRIOR FILING DATE: 2003-05-29
; PRIOR APPLICATION NUMBER: JP 2002-203269
; PRIOR FILING DATE: 2002-05-30
; PRIOR APPLICATION NUMBER: JP 2002-383870
; PRIOR FILING DATE: 2002-12-11
; NUMBER OF SEQ ID NOS: 56791
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 47226
; LENGTH: 548
; TYPE: PRT
; ORGANISM: Oryza sativa
US-10-449-902-47226

Query Match      18.5%; Score 637.5; DB 6; Length 548;
Best Local Similarity 37.9%; Pred. No. 2.3e-26;
Matches 151; Conservative 62; Mismatches 124; Indels 61; Gaps 11;

QY      4 YDELLKYYELHETIGTGGFAKVLAACHILTGMVAIKMDKNTLGS--DLPRIKTEIAL 61
Db      53 HSEALKNYNLGRTLGIGSGFKVIAEHLKTHRVAIKILNRQRNMNEMEEKAKREIKIL 112

QY      62 KNLRHQHICOLYHVLETANKIFMWLEYCPGGELFDYIISQDRLSBEETRVFRQIVSAVA 121
Db      113 RLFPHIIRLYEVIYTPTDIYVMVEYCKFGELFDYIVEKGRLODEARRIFQOIISGYE 172

QY      122 YVHSQGYAHRDLKPENLLFDEYHKLIIDFGLCAKPKGNKDYHLOTCCGSLAYAAPELIQ 181
Db      173 YCHRNMMVVRDLKPENLLDLSKYNVKLADPGL--SNVMHDGHFLKTSCTGSPNYAAPEVIS 230

QY      182 GKSYLEGSEADVWMSGILLYVLMCGFLPFDDNNVMALYKKIMRGKYDVPKWLSPSSILLIQ 241
Db      231 GKLYAGPEVDVWMSGVILYALLCGTLPPDDENIPNLFKKIKGGIYTLPSHLSALARDLIP 290

QY      242 QMLQVDPKGRISMKNNLNHPIMQDYNYP-----VEWQSKNPFHLLDDCVTELSV 292
Db      291 RMLVVDPMKRITIREHQWF--QIRLPYLAVPPTAQOAK-----MIDED----- 337
```



**This Page Blank (uspto)**